

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 02/086443 A2

(51) International Patent Classification⁷: **G01N**

(21) International Application Number: **PCT/US02/12476**

(22) International Filing Date: **18 April 2002 (18.04.2002)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
60/284,770 18 April 2001 (18.04.2001) US
60/290,492 10 May 2001 (10.05.2001) US
60/339,245 9 November 2001 (09.11.2001) US
60/350,666 13 November 2001 (13.11.2001) US
60/334,370 29 November 2001 (29.11.2001) US
60/372,246 12 April 2002 (12.04.2002) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv)) for US only

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(71) Applicant (*for all designated States except US*): **EOS BIOTECHNOLOGY, INC.** [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **AZIZ, Natasha** [US/US]; 411 California Avenue, Palo Alto, CA 94306 (US). **MURRAY, Richard** [US/US]; 22643 Woodbridge Court, Cupertino, CA 95014 (US).

(74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

WO 02/086443 A2

(54) Title: METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

(57) Abstract: Described herein are methods and compositions that can be used for diagnosis and treatment of lung cancer and similar pathologies. Also described herein are methods that can be used to identify modulators of lung cancer and similar pathologies.

METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS
OF SCREENING FOR MODULATORS OF LUNG CANCER

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application is related to USSN 60/284,770, filed April 18, 2001; USSN 60/290,492, filed May 10, 2001; USSN 60/334,370, filed November 29, 2001; USSN 60/339,245, filed November 9, 2001; USSN 60/350,666, filed November 13, 2001; and
10 USSN 60/xxx,xxx, filed April 12, 2002 (Docket OMNI-002P); each of which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression
15 profiles and nucleic acids, products, and antibodies thereto that are involved in lung cancer; and to the use of such expression profiles and compositions in diagnosis and therapy of lung cancer. The invention further relates to methods for identifying and using agents and/or targets that inhibit lung cancer or related conditions.

20

BACKGROUND OF THE INVENTION

Lung cancer is the second most commonly occurring cancer in the United States and is the leading cause of cancer-related death. It is estimated that there are over 160,000 new cases of lung cancer in the United States every year. Of those who are diagnosed with lung cancer, 86 percent will die within five years. Lung cancer is the most common visceral
25 cancer in men and accounts for nearly one third of all cancer deaths in both men and women. In fact, lung cancer accounts for 7% of all deaths, due to any cause, in both men and women.

Smoking is the primary cause of lung cancer, with more than 80% of lung cancers resulting from smoking. About 400 to 500 separate gaseous substances are present in the smoke of a non-filter cigarette. The most noteworthy substances include nitrogen oxides,
30 hydrogen cyanide, formaldehyde, benzene, and toluene. The particles present in cigarette smoke contain at least 3,500 individual compounds such as nicotine, tobacco alkaloids (nornicotine, anatabine, anabasine), polycyclic aromatic hydrocarbons (e.g., benzo(a)pyrene, B(a)P), naphthalenes, aromatic amines, phenols, and tobacco-specific nitrosamines.

Tobacco-specific nitrosamines are formed during tobacco curing and processing, and are suspected of causing lung cancer in humans. In rodent studies, regardless of the where or how it is applied, the tobacco-specific nitrosamine known as NNK produces lung adenomas and lung adenocarcinomas. The tobacco-specific nitrosamine known as NNAL also produces
5 lung adenocarcinomas in rodents.

Many of the chemicals found in cigarette smoke also affect the nonsmoker inhaling "secondhand" or sidestream smoke. Indeed, the smoke inhaled by non-smokers has a chemical composition similar to the smoke inhaled by smokers, but, importantly, the concentrations of the carcinogenic tobacco-specific nitrosamines are present in higher
10 concentrations in second hand smoke. For this and other reasons, "passive smoking" is an important cause of lung cancer, causing as many as 3,000 lung cancer deaths in nonsmokers each year.

In addition to smoking, other factors thought to be causes of lung cancer include on-the-job exposure to carcinogens such as asbestos and uranium, exposure to chemical hazards
15 such as radon, polycyclic aromatic hydrocarbons, chromium, nickel, and inorganic arsenic, genetic factors, and diet.

Histological classification of various lung cancers define the types of cancer that begin in the lung. See, e.g., Travis, et al. (1999) Histological Typing of Lung and Pleural Tumours (International Histological Classification of Tumours, No 1. Four major cell types
20 make up more than 88% of all primary lung neoplasms. These are: squamous or epidermoid carcinoma, small cell (also called oat cell) carcinoma, adenocarcinoma, and large cell (also called large cell anaplastic) carcinoma. The remainder include undifferentiated carcinomas, carcinoids, bronchial gland tumors, and other rarer types. The various cell types have different natural histories and responses to therapy, and, thus, a correct histologic diagnosis is
25 the first step of effective treatment.

Small cell lung cancer (SCLC) accounts for 18-25% of all lung cancers, and occurs less frequently than non-small cell lung cancers, and generally spread to distant organs more rapidly than non-small cell lung cancer. In general, at the time of presentation small cell lung cancers have already spread beyond the bounds where surgery and curative intent
30 can be undertaken. However, if identified early enough, these cancers are often responsive to chemotherapy and thoracic radiation treatment.

Non-small cell lung cancers (NSCLC) are the more frequently occurring form of lung cancer. They comprise squamous cell carcinoma, adenocarcinoma, and large cell carcinoma

and account for more than 75% of all lung cancers. Non-small cell tumors that are localized at the time of presentation can sometimes be cured with surgery and/or radiotherapy, but usually are not identified until significant metastasis has occurred, which are typically not very responsive to surgical, chemotherapy, or radiation treatment..

5 The screening of asymptomatic persons at high risk for lung cancer has often proven ineffective. In general, only 5 to 15 percent of lung cancer patients have their disease detected while they are asymptomatic. Of course, early detection and treatment are critical factors in the fight against lung cancer. The average survival rate is 49% for those whose cancer is detected early, before the cancer has spread from the lung. Lung cancer often
10 spreads outside of the lung, and it may have spread to the bones or brain by the time it is diagnosed. While the prognosis may be better for lung cancers that are detected early, because of the lack ofv effective curative treatments, early detection does not necessarily alter the total death rate from lung cancer.

 Thus, methods for diagnosis and prognosis of lung cancer and effective treatment of
15 lung cancer would be desirable. Accordingly, provided herein are methods that can be used in diagnosis and prognosis of lung cancer. Further provided are methods that can be used to screen candidate therapeutic agents for the ability to modulate, e.g., treat, lung cancer. Additionally, provided herein are molecular targets and compositions for therapeutic intervention in lung disease and other metastatic cancers.

20

SUMMARY OF THE INVENTION

 The present invention provides nucleotide sequences of genes that are up- and down-regulated in lung cancer cells. Such genes are useful for diagnostic purposes, and also as targets for screening for therapeutic compounds that modulate lung cancer, such as
25 antibodies. The methods of detecting nucleic acids of the invention or their encoded proteins can be used for a number of purposes. Examples include early detection of lung cancers, monitoring and early detection of relapse following treatment of lung cancers, monitoring response to therapy of lung cancers, determining prognosis of lung cancers, directing therapy of lung cancers, selecting patients for postoperative chemotherapy or radiation therapy,
30 selecting therapy, determining tumor prognosis, treatment, or response to treatment, and early detection of precancerous lesions of the lung. Examples of benign or precancerous lesions include: atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis (HP), interstitial pulmonary fibrosis (IPF), asthma, and

bronchiectasis. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting a lung cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at
5 least 80% identical to a sequence as shown in Tables 1A-16. Alternatively, the sample may be contacted with a specific binding reagent, e.g., antibody.

In one embodiment, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16. In another embodiment, the
10 polynucleotide comprises a sequence as shown in Tables 1A-16.

In one embodiment, the biological sample is a tissue sample, or a body fluid. In another embodiment, the biological sample comprises isolated nucleic acids, e.g., mRNA.

In one embodiment, the polynucleotide is labeled, e.g., with a fluorescent label. In one embodiment, the polynucleotide is immobilized on a solid surface. In one embodiment,
15 the patient is undergoing a therapeutic regimen to treat lung cancer. In another embodiment, the patient is suspected of having lung cancer. In one embodiment, the patient is a primate, e.g., a human.

In one embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80%
20 identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy. Or the sample may be evaluated for protein, e.g., contacting the sample with an antibody.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated transcript to a level of the lung cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic
30 treatment. Or the sample may be evaluated for comparison of protein.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a

biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated antibody in the biological sample by contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated antibody, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated antibody to a level of the lung cancer-associated antibody in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated polypeptide in the biological sample by contacting the biological sample with an antibody, wherein the antibody specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1A-16. In one embodiment, an expression vector or cell comprises the isolated nucleic acid. In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16. In one embodiment, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical. In one embodiment, the antibody is an antibody fragment. In another embodiment, the antibody is humanized.

In one aspect, the present invention provides a method of detecting lung cancer in a patient, the method comprising contacting a biological sample from the patient with an antibody or protein as described herein.

In another aspect, the present invention provides a method of detecting antibodies specific to a lung cancer gene in a patient, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprises a sequence from Tables 1A-16.

5 In another aspect, the present invention provides a method for identifying a compound that modulates a lung cancer-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a lung cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16; and (ii) determining the functional effect of the
10 compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect. In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant. In one
15 embodiment, the functional effect is determined by measuring ligand binding to the polypeptide.

In another aspect, the present invention provides a method of inhibiting proliferation or another critical process of a lung cancer-associated cell to treat lung cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein. In one embodiment, the compound is an
20 antibody.

In another aspect, the present invention provides a drug screening assay comprising the steps of: (i) administering a test compound to a mammal having lung cancer or a cell isolated therefrom; (ii) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables
25 1A-16 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of lung cancer.

In one embodiment, the control is a mammal with lung cancer or a cell therefrom that has not been treated with the test compound. In another embodiment, the control is a normal
30 cell or mammal, or a non-malignant lung disease.

In another aspect, the present invention provides a method for treating a mammal having lung cancer comprising administering a compound identified by the assay described herein.

In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having lung cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

5

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and treatment of lung disease or cancer, as well as methods for screening for compositions which modulate lung cancer. "Treatment, monitoring, detection or modulation of lung disease or cancer" includes treatment, monitoring, detection, or modulation of lung disease in those patients who have lung disease (whether malignant or non-malignant, e.g., emphysema, bronchitis, or fibrosis) as well as patients with lung cancers in which gene expression from a gene in Tables 1A-16 is increased or decreased, indicating that the subject is more likely to have disease. In particular, while these targets are identified primarily from lung cancer samples, these same targets are likely to be similarly found in analyses of other medical conditions. These other conditions may result from similar pathological processes which affect similar tissues, e.g., lung cancer, small cell lung carcinoma (oat cell carcinoma), non-small cell carcinomas (e.g., squamous cell carcinoma, adenocarcinoma, large cell lung carcinoma, carcinoid, granulomatous), fibrosis (idiopathic pulmonary fibrosis (IPF), hypersensitivity pneumonitis (HP), interstitial pneumonitis, nonspecific idiopathic pneumonitis (NSIP)), chronic obstructive pulmonary disease (COPD, e.g., emphysema, chronic bronchitis), asthma, bronchiectasis, and esophageal cancer. See, e.g., the NCI webpage and USSN 60/347,349 and USSN 60/xxx,xxx (docket LFBR-001-1P, filed March 29, 2002), each of which is incorporated herein by reference. The treatment may be of lung cancer or related condition itself, or treatment of metastasis.

25 In particular, identification of markers selectively expressed on these cancers allows for use of that expression in diagnostic, prognostic, or therapeutic methods. As such, the invention defines various compositions, e.g., nucleic acids, polypeptides, antibodies, and small molecule agonists/antagonists, which will be useful to selectively identify those markers. For example, therapeutic methods may take the form of protein therapeutics which use the marker expression for selective localization or modulation of function (for those markers which have a causative disease effect), for vaccines, identification of binding partners, or antagonism, e.g., using antisense or RNAi. The markers may be useful for molecular characterization of subsets of lung diseases, which subsets may actually require

very different treatments. Moreover, the markers may also be important in related diseases to the specific cancers, e.g., which affect similar tissues in non-malignant diseases, or have similar mechanisms of induction/maintenance. Metastatic processes or characteristics may also be targeted. Diagnostic and prognostic uses are made available, e.g., to subset related but distinct diseases, or to determine treatment strategy. The detection methods may be based upon nucleic acid, e.g., PCR or hybridization techniques, or protein, e.g., ELISA, imaging, IHC, etc. The diagnosis may be qualitative or quantitative, and may detect increases or decreases in expression levels.

Tables 1A-16 provide unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in lung cancer samples. The tables also provide an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. In Table 1A, genes marked as "target 1" or "target 2" are particularly useful as therapeutic targets. Genes marked as "target 3" are particularly useful as diagnostic markers. Genes marked as "chron" are upregulated in chronically diseased lung (e.g., emphysema, bronchitis, fibrosis) relative to lung tumors and normal tissue. In certain analyses, the ratio for the "chron" category was determined using the 70th percentile of chronically diseased lung samples divided by the 90th percentile of normal lung samples. The ratio for the targets was determined using the 70th percentile of lung tumor samples divided by the 90th percentile of normal lung samples.

Definitions

The term "lung cancer protein" or "lung cancer polynucleotide" or "lung cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1A-16 and conservatively modified variants thereof; or (4)

have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "lung cancer polypeptide" and a "lung cancer polynucleotide," include both naturally occurring or recombinant forms.

A "full length" lung cancer protein or nucleic acid refers to a lung cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains the elements normally contained in one or more naturally occurring, wild type lung cancer polynucleotide or polypeptide sequences. The "full length" may be prior to, or after, various stages of post-translational processing or splicing, including alternative splicing.

"Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a lung cancer protein, polynucleotide, or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes, archival materials, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or other mammal; or a bird; reptile; fish. Livestock and domestic animals are of interest.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention in vivo. Archival tissues or materials, having treatment or outcome history, will be particularly useful.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the

same or have a specified percentage of amino acid residues or nucleotides that are the same (e.g., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using, e.g., a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or insertions, substitutions, and naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer

Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel, et al. (eds. 1995 and supplements) Current Protocols in Molecular Biology.

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul, et al. (1977) Nuc. Acids Res. 25:3389-3402 and Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5877). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between

two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be negative large numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture Collection catalog or web site, www.atcc.org).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least about 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant or component from the composition to be purified.

In this sense, purification does not require that the purified compound be homogeneous, e.g., 100% pure.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refer to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain some basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refer to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function similarly to another amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG, and GCU each encode the amino acid alanine. Thus, at each position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of

conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. In certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally similar molecule. Accordingly, a silent variation of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not necessarily with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. Typically conservative substitutions include for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, Proteins (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts, et al. (1994) Molecular Biology of the Cell (3rd ed.) and Cantor and Schimmel (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules. "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

"Nucleic acid" or "oligonucleotide" or "polynucleotide" or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have at least one different linkage, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, in Sanghui and Cook, eds. Carbohydrate Modifications in Antisense Research, ASC Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (T_m) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4° C drop in T_m for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9° C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary

strand; thus the sequences described herein also provide the complement of the sequence.

The nucleic acid may be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine

5 hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g., the individual units of a peptide nucleic
10 acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, physiological, chemical, or other physical means. For example, useful labels include ^{32}P , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins
15 or other entities which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The labels may be incorporated into the cancer nucleic acids, proteins, and antibodies. Many methods known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter, et al. (1962) Nature 144:945; David, et al. (1974) Biochemistry
20 13:1014-1021; Pain, et al. (1981) J. Immunol. Meth., 40:219-230; and Nygren (1982) J. Histochem. and Cytochem. 30:407-412.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody.
25 The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either
30 covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method

using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, e.g., through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, preferably one that does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled, e.g., with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled, e.g., with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered

recombinant for the purposes of the invention. Similarly, a "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is typically an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, e.g., wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed in operable linkage to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule selectively to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to essentially no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in "Overview of principles of hybridization and the strategy of nucleic acid assays" in Tijssen (1993) Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes (vol. 24) Elsevier. Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is typically at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions are often: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C. For PCR, a temperature of about 36° C is typical for low stringency amplification, although annealing temperatures may vary between about 32° C and 48° C depending on primer length. For high stringency PCR amplification, a temperature of about 62° C is typical, although high stringency annealing temperatures can range from about 50° C to about 65° C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90° C - 95° C for 0.5 - 2 min., an annealing phase lasting 0.5 - 2 min., and an extension phase of about 72° C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis, et al.(1990) PCR Protocols, A Guide to Methods and Applications.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This

occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C, and a wash in 1X SSC at 45° C. A positive hybridization is at least twice background. Alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., Ausubel, et al. (ed.) Current Protocols in Molecular Biology Lippincott.

10 The phrase "functional effects" in the context of assays for testing compounds that modulate activity of a lung cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the lung cancer protein or nucleic acid, e.g., a physiological, enzymatic, functional, physical, or chemical effect, such as the ability to decrease lung cancer. It includes ligand binding activity; cell viability, cell growth on soft
15 agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein expression in cells undergoing metastasis, and other characteristics of lung cancer cells. "Functional effects" include *in vitro*, *in vivo*, and *ex vivo* activities.

20 By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a lung cancer protein sequence, e.g., physiological, functional, enzymatic, physical, or chemical effects. Such functional effects can be measured by many means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance,
25 refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the lung cancer protein; measuring binding activity or binding assays, e.g., binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on lung cancer can also be performed using lung cancer assays known to those of
30 skill in the art such as an *in vitro* assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein

expression in cells undergoing metastasis, and other characteristics of lung cancer cells. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for lung cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β -gal, GFP, and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

“Inhibitors”, “activators”, and “modulators” of lung cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using *in vitro* and *in vivo* assays of lung cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of lung cancer proteins, e.g., antagonists. Antisense or inhibitory nucleic acids may seem to inhibit expression and subsequent function of the protein.

“Activators” are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate lung cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of lung cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the lung cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of lung cancer can also be identified by incubating lung cancer cells with the test compound and determining increases or decreases in the expression of 1 or more lung cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more lung cancer proteins, such as lung cancer proteins encoded by the sequences set out in Tables 1A-16.

Samples or assays comprising lung cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of a lung cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more

preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase "changes in cell growth" refers to any change in cell growth and proliferation characteristics *in vitro* or *in vivo*, such as cell viability, formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. See, e.g., Freshney (1994) Culture of Animal Cells a Manual of Basic Technique pp. 231-241 (3rd ed.).

"Tumor cell" refers to precancerous, cancerous, and normal cells in a tumor.

"Cancer cells," "transformed" cells, or "transformation" in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy (see, Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3rd ed.)).

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD, and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul, Fundamental Immunology.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible

for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H - C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Paul (ed. 1999) Fundamental Immunology (4th ed.). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty, et al. (1990) Nature 348:552-554).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler and Milstein (1975) Nature 256:495-497; Kozbor, et al. (1983) Immunology Today 4:72; Cole, et al. (1985), pp. 77-96 in Monoclonal Antibodies and Cancer Therapy; Coligan (1991 and supplements) Current Protocols in Immunology; Harlow and Lane (1988) Antibodies, A Laboratory Manual; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty, et al. (1990) Nature 348:552-554; Marks, et al. (1992) Biotechnology 10:779-783).

A "chimeric antibody" is an antibody molecule in which, e.g., (a) the constant region, or a portion thereof, is altered, replaced, or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function, and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the

variable region, or a portion thereof, is altered, replaced, or exchanged with a variable region having a different or altered antigen specificity.

Identification of lung cancer-associated sequences

5 In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is
10 characteristic of the state of the cell. That is, normal tissue may be distinguished from cancerous or metastatic cancerous tissue, or metastatic cancerous tissue can be compared with tissue from surviving cancer patients. By comparing expression profiles of tissue in known different lung cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.
15 Molecular profiling may distinguish subtypes of a currently collective disease designation, e.g., different forms of lung cancer (chronic disease, adenocarcinoma, etc.)

 The identification of sequences that are differentially expressed in lung cancer versus non-lung cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-
20 regulate lung cancer, and thus tumor growth or recurrence, in a particular patient. Alternatively, a treatment step may induce other markers which may be used as targets to destroy tumor cells. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Malignant disease may be compared to non-malignant conditions. Metastatic tissue can also be analyzed to determine
25 the stage of lung cancer in the tissue, or origin of primary tumor, e.g., metastasis from a remote primary site. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the lung cancer expression profile. This may be done by making biochips comprising sets of the important lung cancer genes,
30 which can then be used in these screens. PCR methods may be applied with selected primer pairs, and analysis may be of RNA or of genomic sequences. These methods can also be done on the protein basis; that is, protein expression levels of the lung cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the lung cancer

nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the lung cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs or as protein or DNA vaccines.

5 Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in lung cancer relative to normal tissues and/or non-malignant lung disease, or in different types of lung disease, herein termed "lung cancer sequences." As outlined below, lung cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in lung cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the lung cancer sequences are from humans; however, as will be appreciated by those in the art, lung cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other lung cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets (dogs, cats, etc.). Lung cancer sequences from other organisms may be obtained using the techniques outlined below.

 Lung cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, lung cancer nucleic acid sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the lung cancer sequences can be generated.

 A lung cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, e.g., using homology programs or hybridization conditions.

 For identifying lung cancer-associated sequences, the lung cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, cancer and non-malignant conditions, non-malignant conditions and normal tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing lung cancer samples with metastatic cancer samples from other cancers, such as, breast, other gastrointestinal cancers, prostate, ovarian,

etc. Samples of, non metastatic disease tissue and tissue undergoing metastasis are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, e.g., from Affymetrix, Santa Clara, CA. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal lung, but also including, and not limited to colon, heart, brain, liver, breast, kidney, muscle, prostate, small intestine, large intestine, spleen, bone, and/or placenta. In a preferred embodiment, those genes identified during the lung cancer screen that are expressed in significant amounts in other tissues (e.g., essential organs) are removed from the profile, although in some embodiments, this is not necessary (e.g., where organs may be dispensable at a later stage of life). That is, when screening for drugs, it is usually preferable that the target expression be disease specific, to minimize possible side effects on other organs.

In a preferred embodiment, lung cancer sequences are those that are up-regulated in lung cancer; that is, the expression of these genes is higher in cancerous tissue than in normal lung or other tissue. "Up-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater. Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. Unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, e.g., Benson, DA, et al (1998) *Nucleic Acids Research* 26:1-7 and <http://www.ncbi.nlm.nih.gov/>. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ). Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. In some situations, the sequences may be derived from assembly of available sequences or be predicted from genomic DNA using exon prediction algorithms, such as FGENESH (Salamov and Solovyev (2000) *Genome Res.* 10:516-522). In other situations, sequences have been derived from cloning and sequencing of isolated nucleic acids.

In another preferred embodiment, lung cancer sequences are those that are down-regulated in the lung cancer; that is, the expression of these genes is lower in cancerous tissue

or normal lung or other tissue. "Down-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater, or, when the ratio is presented as a number less than one, that the ratio is less than one, preferably 0.5 or less, more preferably 0.25 or less.

5

Informatics

The ability to identify genes that are over or under expressed in lung cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with lung cancer. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (see Anderson (1998) Pharmaceutical Proteomics: Targets, Mechanism, and Function, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in a form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample representing lung cancer, i.e., the identification of lung cancer-associated sequences described herein, provide an abundance of information, which can be correlated with

pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

See also Mount, et al. (2001) Bioinformatics; Durbin, et al. (eds., 1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids (; Baxevanis and Ouellette (eds., 1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins); Rashidi and Buehler (1999) Bioinformatics: Basic Applications in Biological Science and Medicine; Setubal, et al. (eds 1997) Introduction to Computational Molecular Biology; Misener and Krawetz (eds, 2000) Bioinformatics: Methods and Protocols; Higgins and Taylor (eds., 2000) Bioinformatics: Sequence, Structure, and Databanks: A Practical

Approach; Brown (2001) Bioinformatics: A Biologist's Guide to Biocomputing and the Internet; Han and Kamber (2000) Data Mining: Concepts and Techniques (2000); and Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes.

The present invention provides a computer database comprising a computer and
5 software for storing in computer-retrievable form assay data records cross-tabulated, e.g.,
with data specifying the source of the target-containing sample from which each sequence
specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample
is from a control tissue sample known to be free of pathological disorders. In a variation, at
10 least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or
another tissue specimen to be analyzed for lung cancer. In another variation, the assay
records cross-tabulate one or more of the following parameters for each target species in a
sample: (1) a unique identification code, which can include, e.g., a target molecular structure
and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample
15 source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data
in a computer data storage apparatus, which can include magnetic disks, optical disks,
magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic
bubble memory devices, and other data storage devices, including CPU registers and on-CPU
20 data storage arrays. Typically, the target data records are stored as a bit pattern in an array of
magnetic domains on a magnetizable medium or as an array of charge states or transistor gate
states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor
and a charge storage area, which may be on the transistor). In one embodiment, the invention
provides such storage devices, and computer systems built therewith, comprising a bit pattern
25 encoding a protein expression fingerprint record comprising unique identifiers for at least 10
target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a
method for identifying related peptide or nucleic acid sequences, comprising performing a
computerized comparison between a peptide or nucleic acid sequence assay record stored in
30 or retrieved from a computer storage device or database and at least one other sequence. The
comparison can include a sequence analysis or comparison algorithm or computer program
embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may

be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, 5 SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices 10 linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

15 The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

20 In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results.

25 Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary 30 memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC,

MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

15 **Characteristics of lung cancer-associated proteins**

Lung cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the lung cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or dysregulated cellular processes (see, e.g., Alberts (ed. 1994) Molecular Biology of the Cell (3d ed.)). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more structural motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2

domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of amino acid sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (see, e.g., Bateman, et al (2000) *Nuc. Acids Res.* 28:263-266; Sonnhammer, et al. (1997) *Proteins* 28:405-420; Bateman, et al. (1999) *Nuc. Acids Res.* 27:260-262; and Sonnhammer, et al. (1998) *Nuc. Acids Res.* 26:320-322).

In another embodiment, the lung cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels, pumps, and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 17 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the

localization and number of transmembrane domains within the protein may be predicted (see, e.g., PSORT web site <http://psort.nibb.ac.jp/>).

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF, and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, hormones, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they may mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains may also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Lung cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for extracellular immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ* or in histological analysis. Alternatively, antibodies can also label intracellular proteins, in which case analytical samples are typically permeabilized to provide access to intracellular proteins. In addition, some membrane proteins can be processed to release a soluble protein, or to expose a residual fragment. Released soluble proteins may be useful diagnostic markers, processed residual protein fragments may be useful lung markers of disease.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the lung cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins may have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; e.g., if circulating, they often serve to transmit

signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor), an endocrine manner (acting on cells at a distance, e.g., secretion into the blood stream), or exocrine (secretion, e.g., through a duct or to adjacent epithelial surface as sweat glands, sebaceous glands, pancreatic ducts, lacrimal glands, mammary glands, wax producing glands of the ear, etc.). Thus secreted molecules often find use in modulating or altering numerous aspects of physiology. Lung cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood, plasma, serum, or stool tests. Those which are enzymes may be antibody or small molecule targets. Others may be useful as vaccine targets, e.g., via CTL mechanisms.

Use of lung cancer nucleic acids

As described above, lung cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

The lung cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1A-16, can be fragments of larger genes, i.e., they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the lung cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, et al., *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once a lung cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire lung cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant lung cancer nucleic acid can be further-used as a probe to identify and isolate

other lung cancer nucleic acids, e.g., extended coding regions. It can also be used as a “precursor” nucleic acid to make modified or variant lung cancer nucleic acids and proteins.

The lung cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the lung cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, RNAi, vaccine, and/or antisense applications. Alternatively, the lung cancer nucleic acids that include coding regions of lung cancer proteins can be put into expression vectors for the expression of lung cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to lung cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the lung cancer nucleic acids, i.e., the target sequence (either the target sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by “substantially complementary” herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under appropriate reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally complements of ORFs or whole genes are not used. In some embodiments, nucleic acids of lengths up to hundreds of bases can be used.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a

particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical
5 equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is typically meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is
10 the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can
15 be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to a biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be
20 synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified for the attachment or association of the nucleic acid probes and is amenable to at least one detection
25 method. Often the substrate may contain discrete individual sites appropriate for individual partitioning and identification. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.),
30 polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in US application entitled Reusable Low Fluorescent Plastic Biochip, U.S.

Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g., using linkers as are known in the art; e.g., homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized, and then attached to the surface of the solid support. Either the 5' or 3' terminus may be attached to the solid support, or attachment may be via linkage to an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized *in situ*, using known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affymetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of lung cancer-associated sequences. These assays are typically performed in conjunction with

reverse transcription. In such assays, a lung cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of lung cancer-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, e.g., in Innis, et al. (1990) PCR Protocols, A Guide to Methods and Applications.

In some embodiments, a TaqMan based assay is used to measure expression.

10 TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (see, e.g., literature provided by Perkin-Elmer, e.g., www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4:560, Landegren, et al. (1988) Science 241:1077, and Barringer, et al. (1990) Gene 89:117), transcription amplification (Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173), self-sustained sequence replication (Guatelli, et al. (1990) Proc. Nat. Acad. Sci. USA 87:1874), dot PCR, and linker adapter PCR, etc.

Expression of lung cancer proteins from nucleic acids

25 In a preferred embodiment, lung cancer nucleic acids, e.g., encoding lung cancer proteins, are used to make a variety of expression vectors to express lung cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (see, e.g., Ausubel, *supra*, and Fernandez and Hoeffler (eds 1999) Gene Expression Systems) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the lung cancer protein. The term "control sequences" refers to DNA

sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

5 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding
10 linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in
15 accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the lung cancer protein. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are
20 not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences may be either constitutive or inducible promoters. The promoters
25 may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two
30 organisms, e.g., in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector often contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating

vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez and Hoeffler, *supra*).

5 In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The lung cancer proteins of the present invention are usually produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a lung cancer protein, under the appropriate conditions to induce or cause expression of the lung cancer protein. Conditions appropriate for lung cancer protein expression will vary with the
10 choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate
15 growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae*
20 and other yeasts, *E. coli*, *Bacillus subtilis*, S9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the lung cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and
25 adenoviral systems. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (see, e.g., Fernandez and Hoeffler, *supra*). Typically, transcription termination and
30 polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, lung cancer proteins are expressed in bacterial systems. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the lung cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez and Hoeffler, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, lung cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, lung cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*,

Kluyveromyces fragilis and *K. lactis*, *Pichia guillermondii*, and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

The lung cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the lung cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the lung cancer protein may be made as a fusion protein to increase expression for affinity purification purposes, or for other reasons. For example, when the lung cancer protein is a lung cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

In a preferred embodiment, the lung cancer protein is purified or isolated after expression. Lung cancer proteins may be isolated or purified in a variety of appropriate ways. Standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the lung cancer protein may be purified using a standard anti-lung cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes (1982) Protein Purification. The degree of purification necessary will vary depending on the use of the lung cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the lung cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, therapeutic entities, for production of antibodies, as transcription or translation inhibitors, etc.

Variants of lung cancer proteins

In one embodiment, the lung cancer proteins are derivative or variant lung cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative lung cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at a particular residue within the lung cancer peptide.

Also included within one embodiment of lung cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three

classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the lung cancer protein, using cassette or PCR mutagenesis or other techniques, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above.

5 However, variant lung cancer protein fragments having up to about 100-150 residues may be prepared by *in vitro* synthesis. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the lung cancer protein amino acid sequence. The variants typically exhibit a similar qualitative biological activity as the naturally occurring analogue,
10 although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is often predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be
15 conducted at the target codon or region and the expressed lung cancer variants screened for the optimal combination of desired activity. Techniques exist for making substitution mutations at predetermined sites in DNA having a known sequence, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of mutants is often done using assays of lung cancer protein activities.

20 Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be occasionally tolerated. Deletions generally range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive
25 at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. Larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of a lung cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution chart provided in the definition section.

30 Variants typically exhibit essentially the same qualitative biological activity and will elicit the same immune response as a naturally-occurring analog, although variants also are selected to modify the characteristics of lung cancer proteins as needed. Alternatively, the

variant may be designed or reorganized such that a biological activity of the lung cancer protein is altered. For example, glycosylation sites may be added, altered, or removed.

Covalent modifications of lung cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a lung cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a lung cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking lung cancer polypeptides to a water-insoluble support matrix or surface for use in a method for purifying anti-lung cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimide.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of serinyl, threonyl or tyrosyl residues, methylation of the γ -amino groups of lysine, arginine, and histidine side chains (Creighton (1983) Proteins: Structure and Molecular Properties, pp. 79-86), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the lung cancer polypeptide encompassed by this invention is an altered native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended herein to mean adding to or deleting one or more carbohydrate moieties of a native sequence lung cancer polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express lung cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to lung cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native sequence lung cancer polypeptide (for O-linked glycosylation sites). The lung cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the lung cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the lung cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and in Aplin and Wriston (1981) CRC Crit. Rev. Biochem., pp. 259-306.

5 Removal of carbohydrate moieties present on the lung cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al. (1987) Arch. Biochem. Biophys., 259:52 and by Edge, et al. (1981) Anal. Biochem., 118:131.

10 Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura, et al. (1987) Meth. Enzymol., 138:350.

Another type of covalent modification of lung cancer comprises linking the lung cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent
15 Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192, or 4,179,337.

Lung cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a lung cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric
20 molecule comprises a fusion of a lung cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the lung cancer polypeptide. The presence of such epitope-tagged forms of a lung cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the lung
25 cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a lung cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

30 Various tag polypeptides and their respective antibodies are well known and examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field, et al. (1988) Mol. Cell. Biol. 8:2159-2165); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies

thereto (Evan, et al. (1985) Molecular and Cellular Biology 5:3610-3616); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky, et al. (1990) Protein Engineering 3(6):547-553). Other tag polypeptides include the Flag-peptide (Hopp, et al. (1988) BioTechnology 6:1204-1210); the KT3 epitope peptide (Martin, et al. (1992) Science 255:192-194); tubulin epitope peptide (Skinner, et al. (1991) J. Biol. Chem. 266:15163-15166); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:6393-6397).

Also included are other lung cancer proteins of the lung cancer family, and lung cancer proteins from other organisms, which are cloned and expressed as outlined below.

Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related lung cancer proteins from primates or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include unique areas of the lung cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. PCR reaction conditions are well known in the art (e.g., Innis, PCR Protocols, *supra*).

Antibodies to lung cancer proteins

In a preferred embodiment, when a lung cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the lung cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller lung cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are well known (e.g., Coligan, *supra*; and Harlow and Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of Tables 1A-16 or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal

being immunized. Immunogenic proteins include, e.g., keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Adjuvants include, e.g., Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein (1975) Nature 256:495. In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of the tables, or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding (1986) Monoclonal Antibodies: Principles and Practice, pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovin, or primate origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are typically monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid of the tables or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to lung cancer protein are capable of reducing or eliminating a biological function of a lung cancer protein, in a naked form or conjugated to an effector moiety. That is, the addition of anti-lung cancer protein antibodies (either polyclonal or preferably monoclonal) to lung cancer tissue (or cells containing lung cancer) may reduce or eliminate the lung cancer. Generally, at least a 25% decrease in activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the lung cancer proteins are humanized antibodies (e.g., Xenerex Biosciences, Medarex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of a human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. A humanized antibody optimally also will typically comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-329; and Presta (1992) Curr. Op. Struct. Biol. 2:593-596). Humanization can be performed following the method of Winter and co-workers (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-327; Verhoeven, et al. (1988) Science 239:1534-1536), by substituting rodent CDRs or CDR sequences for corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by corresponding sequence from a non-human species.

Human-like antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom and Winter (1991) J. Mol. Biol. 227:381; Marks, et al. (1991) J. Mol. Biol. 222:581). The techniques of Cole, et al. and Boerner, et al. are also available for the preparation of human monoclonal antibodies (Cole, et al. (1985) Monoclonal Antibodies and Cancer Therapy, p. 77 and Boerner, et al. (1991) J. Immunol. 147(1):86-95). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in nearly all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks, et al. (1992) Bio/Technology 10:779-783; Lonberg, et al. (1994) Nature 368:856-859; Morrison (1994) Nature 368:812-13; Fishwild, et al. (1996) Nature Biotechnology 14:845-51; Neuberger (1996) Nature Biotechnology 14:826; and Lonberg and Huszar (1995) Intern. Rev. Immunol. 13:65-93.

By immunotherapy is meant treatment of lung cancer with an antibody raised against a lung cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. The antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the lung cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, may bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted lung cancer protein.

In another preferred embodiment, the lung cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment may bind the extracellular domain of the lung cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The

antibody may cause down-regulation of the transmembrane lung cancer protein. The antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the lung cancer protein. The antibody may be an antagonist of the lung cancer protein or may prevent activation of a transmembrane lung cancer protein, or
5 may induce or suppress a particular cellular pathway. In some embodiments, when the antibody prevents the binding of other molecules to the lung cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF- α , TNF- β , IL-1, INF- γ , and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate,
10 and the like. In some instances the antibody may belong to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, lung cancer may be treated by administering to a patient antibodies directed against the transmembrane lung cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide
15 means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be various molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of a lung cancer protein. In another
20 aspect the therapeutic moiety may modulate an activity of molecules associated with or in close proximity to a lung cancer protein. The therapeutic moiety may inhibit enzymatic or signaling activity such as protease or collagenase activity associated with lung cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to lung cancer tissue or cells results in a reduction
25 in the number of afflicted cells, thereby reducing symptoms associated with lung cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin, saporin, auristatin, and the like. Cytotoxic agents also include
30 radiochemicals made by conjugating radioisotopes to antibodies raised against lung cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane lung cancer proteins not only serves to increase the local concentration of therapeutic moiety in the lung cancer

afflicted area, but also serves to reduce deleterious side effects that may be associated with the untargeted therapeutic moiety.

In another preferred embodiment, the lung cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein or other entity which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the lung cancer protein can be targeted within a cell, i.e., the nucleus, an antibody theretomay contain a signal for that target localization, i.e., a nuclear localization signal.

The lung cancer antibodies of the invention specifically bind to lung cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a K_d of at least about 0.1 mM, more usually at least about 1 μ M, preferably at least about 0.1 μ M or better, and most preferably, 0.01 μ M or better. Selectivity of binding to the specific target and not to related other sequences is also important.

Detection of lung cancer sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the lung cancer phenotype. Expression levels of genes in normal tissue (e.g., not undergoing lung cancer), in lung cancer tissue (and in some cases, for varying severities of lung cancer that relate to prognosis, as outlined below), or in non-malignant disease are evaluated to provide expression profiles. A gene expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state of the cell. While two states may have a particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g.,

normal versus lung cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart (1996) Nature Biotechnology 14:1675-1680, hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is typically at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the RNA or DNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the lung cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to lung cancer genes, e.g., those identified as being important in a lung cancer or disease phenotype, can be evaluated in a lung cancer diagnostic test. In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes.

The lung cancer nucleic acid probes may be attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In a preferred embodiment nucleic acids encoding the lung cancer protein are detected. Although DNA or RNA encoding the lung cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a lung cancer protein is detected.

Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed *in situ*. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a lung cancer protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, lung cancer proteins, including intracellular, transmembrane, or secreted proteins, find use as markers of lung cancer, e.g., for prognostic or diagnostic purposes. Detection of these proteins in putative lung cancer tissue allows for detection, prognosis, or diagnosis of lung cancer or similar disease, and perhaps for selection of therapeutic strategy. In one embodiment, antibodies are used to detect lung cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the lung cancer protein is detected, e.g., by immunoblotting with antibodies raised against the lung cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the lung cancer protein find use in *in situ* imaging techniques, e.g., in histology (e.g., Asai (ed. 1993) Methods in Cell Biology:

Antibodies in Cell Biology, volume 37. In this method cells are contacted with from one to many antibodies to the lung cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label, e.g., multicolor fluorescence or confocal imaging. In another method the primary antibody to the lung cancer protein(s) contains a detectable label, e.g., an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of lung cancer proteins. Many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing lung cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of lung cancer proteins. Antibodies can be used to detect a lung cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous lung cancer protein or vaccine.

In a preferred embodiment, *in situ* hybridization of labeled lung cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including lung cancer tissue and/or normal tissue, are made. *In situ* hybridization (see, e.g., Ausubel, *supra*) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to lung cancer, clinical, pathological, or other information, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. Single or multiple

genes may be useful in various combinations. As above, lung cancer probes may be attached to biochips for the detection and quantification of lung cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

5

Assays for therapeutic compounds

In a preferred embodiment, the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in drug screening
10 assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, et al. (1998) Science 279:84-8; Heid (1996) Genome Res. 6:986-94.

15 In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified lung cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the lung cancer phenotype or an identified physiological function of a lung cancer protein. As above, this can be done on an individual gene level or
20 by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

Having identified differentially expressed genes herein, a variety of assays may be
25 performed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene with altered regulation in lung cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the lung cancer protein. "Modulation" thus includes an increase or a decrease in gene expression. The preferred amount of modulation will depend on the original change of the
30 gene expression in normal versus tissue undergoing lung cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in lung cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in lung

cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the lung cancer protein and standard immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene or protein expression monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

In this embodiment, the lung cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more lung cancer-associated sequences, e.g., a polynucleotide sequence set out in the tables. Generally, in a preferred embodiment, a test compound is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate lung cancer, modulate lung cancer proteins, bind to a lung cancer protein, or interfere with the binding of a lung cancer protein and an antibody, substrate, or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes a molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the lung cancer phenotype or the expression of a lung cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles of nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a lung cancer phenotype, e.g., to a normal or non-malignant tissue fingerprint. In another embodiment, a modulator induces a lung cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

In one aspect, a modulator will neutralize the effect of a lung cancer protein. By “neutralize” is meant that activity of a protein and the consequent effect on the cell is inhibited or blocked.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a lung cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a “lead compound”) with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such “combinatorial chemical libraries” are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop, et al. (1994) J. Med. Chem. 37(9):1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) Pept. Prot. Res. 37:487-493, Houghton, et al. (1991) Nature, 354:84-88), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs, et al. (1993) Proc. Nat. Acad. Sci. USA 90:6909-6913), vinylogous polypeptides (Hagihara, et al. (1992) J. Amer. Chem. Soc. 114:6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, et

al. (1992) J. Amer. Chem. Soc. 114:9217-9218), analogous organic syntheses of small compound libraries (Chen, et al. (1994) J. Amer. Chem. Soc. 116:2661), oligocarbamates (Cho, et al. (1993) Science 261:1303), and/or peptidyl phosphonates (Campbell, et al. (1994) J. Org. Chem. 59:658). See, generally, Gordon, et al. (1994) J. Med. Chem. 37:1385, nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn, et al. (1996) Nature Biotechnology 14(3):309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang, et al. (1996) Science 274:1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. The above devices, with appropriate modification, are suitable for use with the present invention. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, RU, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect modulation of lung cancer gene transcription, polypeptide expression, and polypeptide activity.

High throughput assays for evaluating the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins,

U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g.,
5 Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate procedures, including sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a
10 high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or
15 fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly
20 useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring
25 proteins, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that the nucleic acid or peptide consists of essentially random sequences of nucleotides and amino acids, respectively. Since these random peptides (or nucleic acids, discussed below) are often chemically synthesized, they may incorporate a nucleotide or amino acid at any position. The synthetic process can be designed to generate
30 randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. In a preferred embodiment, the nucleotides or amino acid residues are
5 randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc.

Modulators of lung cancer can also be nucleic acids, as defined above.

10 As described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. Digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

15 After a candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence is analyzed. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro* transcription
20 with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as,
25 alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin
30 is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

Nucleic acid assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos.

5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the lung cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state,

screens are performed to identify agents that bind and/or modulate the biological activity of the gene product, or evaluate genetic polymorphisms.

Genes can be screened for those that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a lung cancer expression pattern leading to a normal expression pattern, or to modulate a single lung cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated lung cancer tissue reveals genes that are not expressed in normal tissue or lung cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for lung cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated lung cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of lung cancer cells, that have an associated lung cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once a test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., lung cancer or non-malignant tissue may be screened for agents that modulate, e.g., induce or suppress a lung cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on lung cancer activity. By defining such a signature for the lung cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

Measure of lung cancer polypeptide activity, or of lung cancer or the lung cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the metastatic polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used
5 to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of lung cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as
10 cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian lung cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed *in vitro*. For example, a lung cancer polypeptide is first contacted with a potential modulator and
15 incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the lung cancer polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is typically measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the lung cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using
20 PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is typically detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a lung cancer protein
25 promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or β -gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual
30 genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed

genes are sometimes referred to herein as "lung cancer proteins." The lung cancer protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another
5 embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated
10 gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the lung cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a lung cancer
15 protein and a candidate compound, and determining the binding of the compound to the lung cancer protein. Preferred embodiments utilize the human lung cancer protein, although other mammalian proteins may also be used, e.g., for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative lung cancer proteins may be used.

20 Generally, in a preferred embodiment of the methods herein, the lung cancer protein or the candidate agent is non-diffusably bound to an insoluble support, preferably having isolated sample receiving areas (e.g., a microtiter plate, an array, etc.). The insoluble supports may be made of a composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of
25 screening. The surface of such supports may be solid or porous and of a convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents
30 and samples. The particular manner of binding of the composition is typically not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition, and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation

sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the lung cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the lung cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the lung cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the lung cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g., ^{125}I for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor may be a binding moiety known to bind to the target molecule (i.e., a lung cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically

between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by a test compound. Displacement of the competitor is an indication that the test compound is binding to the lung cancer protein and thus is capable of binding to, and potentially modulating, the activity of the lung cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the lung cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the lung cancer protein.

In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the lung cancer proteins. In one embodiment, the methods comprise combining a lung cancer protein and a competitor in a first sample. A second sample comprises a test compound, a lung cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the lung cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the lung cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native lung cancer protein, but cannot bind to modified lung cancer proteins. The structure of the lung cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a lung cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results.

Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a lung cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising lung cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a lung cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g., hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (e.g., cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate lung cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the lung cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting lung cancer cell division is provided. The method comprises administration of a lung cancer inhibitor. In another embodiment, a method of inhibiting lung cancer is provided. The method may comprise administration of a lung cancer inhibitor. In a further embodiment, methods of treating cells or individuals with lung cancer are provided, e.g., comprising administration of a lung cancer inhibitor.

In one embodiment, a lung cancer inhibitor is an antibody as discussed above. In another embodiment, the lung cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, viability, and metastasis assays are known to those of skill in the art, as described below.

Soft agar growth or colony formation in suspension

- 5 Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and
- 10 grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of lung cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.
- 15 Techniques for soft agar growth or colony formation in suspension assays are described in Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3rd ed.), herein incorporated by reference. See also, the methods section of Garkavtsev, et al. (1996), *supra*, herein incorporated by reference.

20 *Contact inhibition and density limitation of growth*

- Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a
- 25 higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (³H)-thymidine at saturation density can be used to measure density limitation of growth. See Freshney (1994), *supra*. The transformed cells, when transfected with tumor suppressor genes, regenerate a
- 30 normal phenotype and become contact inhibited and would grow to a lower density.

 In this assay, labeling index with (³H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a lung cancer-associated sequence and are grown for 24 hours at saturation density in

non-limiting medium conditions. The percentage of cells labeling with (^3H)-thymidine is determined autoradiographically. See, Freshney (1994), *supra*.

Growth factor or serum dependence

5 Transformed cells typically have a lower serum dependence than their normal counterparts (see, e.g., Temin (1966) J. Natl. Cancer Insti. 37:167-175; Eagle, et al. (1970) J. Exp. Med. 131:836-879); Freshney, *supra*. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

Tumor specific markers levels

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g.,
15 Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman (1992) "Angiogenesis and Cancer" in Sem Cancer Biol.

Various techniques which measure the release of these factors are described in
20 Freshney (1994), *supra*. Also, see, Unkeless, et al. (1974) J. Biol. Chem. 249:4295-4305; Strickland and Beers (1976) J. Biol. Chem. 251:5694-5702; Whur, et al. (1980) Br. J. Cancer 42:305-312; Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184; Freshney Anticancer Res. 5:111-130 (1985).

Invasiveness into Matrigel

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate lung cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and
30 invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with ^{125}I and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), *supra*.

Tumor growth in vivo

Effects of lung cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the lung cancer gene is disrupted or in which a lung cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous lung cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous lung cancer gene with a mutated version of the lung cancer gene, or by mutating the endogenous lung cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi, et al. (1989) Science 244:1288). Chimeric targeted mice can be derived according to Hogan, et al. (1988) Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory and Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Washington, D.C.

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic "nude" mouse (see, e.g., Giovanella, et al. (1974) J. Natl. Cancer Inst. 52:921), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley, et al. (1978) Br. J. Cancer 38:263; Selby, et al. (1980) Br. J. Cancer 41:52) can be used as a host. Transplantable tumor cells (typically about 10^6 cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a lung cancer-associated sequences are injected subcutaneously. After a suitable length of time,

preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

5 Polynucleotide modulators of lung cancer

Antisense and RNAi Polynucleotides

In certain embodiments, the activity of a lung cancer-associated protein is downregulated, or entirely inhibited, by the use of antisense or an inhibitory polynucleotide, i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a
10 coding mRNA nucleic acid sequence, e.g., a lung cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their
15 close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the lung cancer protein mRNA. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

20 Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides.
25 Sense oligonucleotides can, e.g., be employed to block transcription by binding to the antisense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for lung cancer molecules. A preferred antisense molecule is for a lung cancer sequence in the tables, or for a ligand or activator thereof. Antisense or sense
30 oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein

is described in, e.g., Stein and Cohen (1988) Cancer Res. 48:2659 and van der Krol, et al. (1988) BioTechniques 6:958).

RNA interference is a mechanism to suppress gene expression in a sequence specific manner. See, e.g., Brumelkamp, et al. (2002) Scienceexpress (21March2002); Sharp (1999) Genes Dev. 13:139-141; and Cathew (2001) Curr. Op. Cell Biol. 13:244-248. In mammalian cells, short, e.g., 21 nt, double stranded small interfering RNAs (siRNA) have been shown to be effective at inducing an RNAi response. See, e.g., Elbashir, et al. (2001) Nature 411:494-498. The mechanism may be used to downregulate expression levels of identified genes, e.g., treatment of or validation of relevance to disease.

Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of lung cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto, et al. (1994) Adv. in Pharmacology 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel, et al. (1990) Nucl. Acids Res. 18:299-304; European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Yamada, et al. (1994) Human Gene Therapy 1:39-45; Leavitt, et al. (1995) Proc. Natl. Acad. Sci. USA 92:699-703; Leavitt, et al. (19994) Human Gene Therapy 5:1151-120; and Yamada, et al. (1994) Virology 205: 121-126).

Polynucleotide modulators of lung cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of lung cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by

formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Thus, in one embodiment, methods of modulating lung cancer in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-lung cancer antibody that reduces or eliminates the biological activity of an endogenous lung cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a lung cancer protein. This may be accomplished in any number of ways. In a preferred embodiment, e.g., when the lung cancer sequence is down-regulated in lung cancer, such state may be reversed by increasing the amount of lung cancer gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous lung cancer gene or administering a gene encoding the lung cancer sequence, using known gene-therapy techniques. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), e.g., as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, e.g., when the lung cancer sequence is up-regulated in lung cancer, the activity of the endogenous lung cancer gene is decreased, e.g., by the administration of a lung cancer antisense or RNAi nucleic acid.

In one embodiment, the lung cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to lung cancer proteins. Similarly, the lung cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify lung cancer antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred embodiment, the antibodies are generated to epitopes unique to a lung cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. The lung cancer antibodies may be coupled to standard affinity chromatography columns and used to purify lung cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the lung cancer protein.

30 **Methods of identifying variant lung cancer-associated sequences**

Without being bound by theory, expression of various lung cancer sequences is correlated with lung cancer. Accordingly, disorders based on mutant or variant lung cancer genes may be determined. In one embodiment, the invention provides methods for

identifying cells containing variant lung cancer genes, e.g., determining all or part of the sequence of at least one endogenous lung cancer genes in a cell. In a preferred embodiment, the invention provides methods of identifying the lung cancer genotype of an individual, e.g., determining all or part of the sequence of at least one lung cancer gene of the individual.

- 5 This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced lung cancer gene to a known lung cancer gene, i.e., a wild-type gene.

10 The sequence of all or part of the lung cancer gene can then be compared to the sequence of a known lung cancer gene to determine if any differences exist. This can be done using known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the lung cancer gene of the patient and the known lung cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

- 15 In a preferred embodiment, the lung cancer genes are used as probes to determine the number of copies of the lung cancer gene in the genome.

In another preferred embodiment, the lung cancer genes are used as probes to determine the chromosomal localization of the lung cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when
20 chromosomal abnormalities such as translocations, and the like are identified in the lung cancer gene locus.

Administration of pharmaceutical and vaccine compositions

- 25 In one embodiment, a therapeutically effective dose of a lung cancer protein or modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel, et al. (1992) Pharmaceutical Dosage Forms and Drug Delivery; Lieberman, Pharmaceutical Dosage Forms (vols. 1-3), Dekker, ISBN 0824770846,
30 082476918X, 0824712692, 0824716981; Lloyd (1999) The Art, Science and Technology of Pharmaceutical Compounding; and Pickar (1999) Dosage Calculations). Adjustments for lung cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration,

drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art.

A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and
5 veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the lung cancer proteins and modulators thereof of the present invention can be done in a variety of ways, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly,
10 intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the lung cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a lung cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the
15 pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid,
20 sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases
25 such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic
30 ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose,

lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms
5 suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that lung cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, etc.) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by
10 packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a lung cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions
15 are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate
20 and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., Remington's Pharmaceutical Science (15th ed., 1980) and Hardman, et al. (eds. 1996) Goodman and Gilman: The Pharmacological Basis of Therapeutics).

25 Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally
30 administrable compositions will be known or apparent to those skilled in the art, e.g., Remington's Pharmaceutical Science and Goodman and Gilman, The Pharmacological Basis of Therapeutics, *supra*.

The compositions containing modulators of lung cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to
5 accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the
10 patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be
15 used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer based, at least in part, upon gene expression profiles. Vaccine strategies may be used, in either a DNA vaccine form, or protein vaccine.

It will be appreciated that the present lung cancer protein-modulating compounds can
20 be administered alone or in combination with additional lung cancer modulating compounds or with other therapeutic agent, e.g., other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, e.g., polynucleotides comprising nucleic acid sequences set forth in the tables, such as antisense or RNAi polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present
25 invention provides methods, reagents, vectors, and cells useful for expression of lung cancer-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo*, or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for
30 introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see, e.g.,

Berger and Kimmel, Guide to Molecular Cloning Techniques. Methods in Enzymology volume 152 (Berger), Ausubel, et al. (eds. 1999) Current Protocols (supplemented through 1999), and Sambrook, et al. (1989) Molecular Cloning - A Laboratory Manual (2nd ed., Vol. 1-3).

5 In a preferred embodiment, lung cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, lung cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the lung cancer coding regions) can be administered in a gene therapy application. These lung cancer genes can include antisense or inhibitory applications, e.g., as inhibitory RNA or gene
10 therapy (e.g., for incorporation into the genome) or as antisense compositions.

Lung cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL, and antibody responses.. Such vaccine compositions can include, e.g., lipidated peptides (see, e.g., Vitiello, et al. (1995) J. Clin. Invest. 95:341), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres
15 (see, e.g., Eldridge, et al. (1991) Molec. Immunol. 28:287-294; Alonso, et al. (1994) Vaccine 12:299-306; Jones, et al. (1995) Vaccine 13:675-681), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi, et al. (1990) Nature 344:873-875; Hu, et al. (1998) Clin Exp Immunol. 113:235-243), multiple antigen peptide systems (MAPs) (see, e.g., Tam (1988) Proc. Natl. Acad. Sci. U.S.A. 85:5409-5413; Tam
20 (1996) J. Immunol. Methods 196:17-32), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, et al., p. 379 In: Kaufmann (ed. 1996) Concepts in vaccine development; Chakrabarti, et al. (1986) Nature 320:535; Hu, et al. (1986) Nature 320:537; Kieny, et al. (1986) AIDS Bio/Technology 4:790; Top, et al. (1971) J. Infect. Dis. 124:148; Chanda, et al.
25 (1990) Virology 175:535), particles of viral or synthetic origin (see, e.g., Kofler, et al. (1996) J. Immunol. Methods 192:25; Eldridge, et al. (1993) Sem. Hematol. 30:16; Falo, et al. (1995) Nature Med. 7:649), adjuvants (Warren, et al. (1986) Annu. Rev. Immunol. 4:369; Gupta, et al. (1993) Vaccine 11:293), liposomes (Reddy, et al. (1992) J. Immunol. 148:1585; Rock (1996) Immunol. Today 17:131), or, naked or particle absorbed cDNA (Ulmer, et al. (1993) Science 259:1745; Robinson, et al. (1993) Vaccine 11:957; Shiver, et al., p. 423 In:
30 Kaufmann (ed. 1996) Concepts in vaccine development; Cease and Berzofsky (1994) Annu. Rev. Immunol. 12:923 and Eldridge, et al. (1993) Sem. Hematol. 30:16). Toxin-targeted

delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff, et. al. (1990) *Science* 247:1465 as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode lung cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, et al. (1991) *Nature* 351:456-460. A wide variety of other vectors useful for therapeutic administration or immunization e.g., adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the

like, will be apparent to those skilled in the art from the description herein (see, e.g., Shata, et al. (2000) Mol Med Today 6:66-71; Shedlock, et al. (2000) J. Leukoc. Biol. 68:793-806; Hipp, et al. (2000) In Vivo 14:571-85).

Methods for the use of genes as DNA vaccines are well known, and include placing a lung cancer gene or portion of a lung cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a lung cancer patient. The lung cancer gene used for DNA vaccines can encode full-length lung cancer proteins, but more preferably encodes portions of the lung cancer proteins including peptides derived from the lung cancer protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a lung cancer gene. For example, lung cancer-associated genes or sequence encoding subfragments of a lung cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the lung cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

In another preferred embodiment lung cancer genes find use in generating animal models of lung cancer. When the lung cancer gene identified is repressed or diminished in metastatic tissue, gene therapy technology, e.g., wherein antisense or inhibitory RNA directed to the lung cancer gene will also diminish or repress expression of the gene. Animal models of lung cancer find use in screening for modulators of a lung cancer-associated sequence or modulators of lung cancer. Similarly, transgenic animal technology including gene knockout technology, e.g., as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the lung cancer protein. When desired, tissue-specific expression or knockout of the lung cancer protein may be necessary.

It is also possible that the lung cancer protein is overexpressed in lung cancer. As such, transgenic animals can be generated that overexpress the lung cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene.

Animals generated by such methods will find use as animal models of lung cancer and are additionally useful in screening for modulators to treat lung cancer.

Kits for Use in Diagnostic and/or Prognostic Applications

5 For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In diagnostic and research applications such kits may include at least one of the following: assay reagents, buffers, lung cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, RNAi, dominant negative lung cancer polypeptides or polynucleotides, small molecule inhibitors of
10 lung cancer-associated sequences, etc. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

 In addition, the kits may include instructional materials containing instructions (e.g., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. A medium
15 capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

 The present invention also provides for kits for screening for modulators of lung
20 cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a lung cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing lung cancer-associated activity. Optionally, the kit contains biologically active lung cancer protein. A wide variety of kits and components can be prepared according to the present
25 invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes typically will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

EXAMPLES

Example 1: Gene Chip Analysis

Molecular profiles of various normal and cancerous tissues were determined and
5 analyzed using gene chips. RNA was isolated and gene chip analysis was performed as
described (Glynne, et al. (2000) Nature 403:672-676; Zhao, et al. (2000) Genes Dev. 14:981-
993).

Tables 1A and 1B were previously filed on April 18, 2001 in USSN 60/284,770 (18501-001500US) and on November 29, 2001 in USSN 60/334,370 (18501-001520US)

5	Table 1A			70% chron/90% NL	70% SQAD/90% NL
	Pkey	ExAccn	UnigenelD	Unigene Title	
	100134	D13264	Hs.49	macrophage scavenger receptor 1	1.61
	100780	HG3731-HT4001		***immunoglobulin Heavy Chain, VdJrc Reg	2.68
10	100971	J02874	Hs.83213	fatty acid binding protein 4; adipocyte	1.96
	101088	L05568	Hs.553	solute carrier family 6 (neurotransmitter)	0.79
	101102	L07594	Hs.79059	transforming growth factor; beta receptor	2.55
	101168	L15388	Hs.211569	G protein-coupled receptor kinase 5	0.88
	101277	L38486	Hs.118223	microfibrillar-associated protein 4	0.89
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.59
15	101336	L49169	Hs.75678	FBJ murine osteosarcoma viral oncogene h	1.15
	101345	L76380	Hs.152175	calcitonin receptor-like	0.81
	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a I	1.31
	101764	M80563	Hs.81256	S100 calcium-binding protein A4 (calcium	1.44
	101771	M81750	Hs.153837	myeloid cell nuclear differentiation ant	0.96
20	101842	M93221	Hs.75182	mannose receptor; C type 1	1.27
	102283	U31384	Hs.83381	guanine nucleotide binding protein 11	1.04
	102363	U39447	Hs.198241	amine oxidase; copper containing 3 (vasc	0.96
	102507	U52154	Hs.193044	potassium inwardly-rectifying channel; s	2.81
	102698	U75272	Hs.1867	progastricsin (pepsinogen C)	0.95
25	103025	X54131	Hs.123641	protein tyrosine phosphatase; receptor t	1.62
	103280	X79981	Hs.76206	cadherin 5; VE-cadherin (vascular epithe	0.9
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2	1.27
	103541	Z11697	Hs.79197	CD83 antigen (activated B lymphocytes; I	1.86
30	103554	Z18951	Hs.74034	caveolin 1; caveolae protein; 22kD	1.27
	104212	AB002298	Hs.173035	KIAA0300 protein	1.17
	104691	AA011176	Hs.37744	ESTs	1.08
	104825	AA035613	Hs.141883	ESTs	0.75
	104857	AA043219	Hs.19058	ESTs	2.6
35	104865	AA045136	Hs.22575	ESTs	1.23
	104989	AA102098	Hs.118615	ESTs	0.63
	105729	AA292694	Hs.3807	ESTs; Weakly similar to PHOSPHOLEMMAN PR	0.86
	105847	AA398606	Hs.32241	ESTs	1.32
	105894	AA400979	Hs.25691	calcitonin receptor-like receptor activi	0.78
40	106490	AA451861	Hs.115537	ESTs; Weakly similar to dipeptidase prec	1.2
	106536	AA453997	Hs.23804	ESTs	0.82
	106605	AA457718	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	0.99
	106667	AA461086	Hs.16578	ESTs	1.17
	106773	AA478109	Hs.188833	ESTs	1.46
45	106797	AA478962	Hs.169943	ESTs	1.18
	106844	AA485055	Hs.158213	sperm associated antigen 6	0.98
	106870	AA487576	Hs.26530	serum deprivation response (phosphatidyl	1.05
	106954	AA496980	Hs.204038	ESTs	1.25
	107054	AA500150	Hs.14366	ESTs	1.11
50	107292	T30407	Hs.4789	ESTs; Weakly similar to oxidative-stress	1.07
	107994	AA036811	Hs.165030	ESTs	0.7
	107997	AA037388	Hs.82223	Human DNA sequence from clone 141H5 on c	1.02
	108041	AA041552	Hs.61957	ESTs	1.44
	108087	AA045709	Hs.40545	ESTs	1.98
55	108382	AA074885	Hs.67726	macrophage receptor with collagenous str	1.52
	108435	AA078787	Hs.194101	ESTs	2.53
	108480	AA081093	Hs.68055	ESTs	1.56
	109252	AA194830	Hs.85944	ESTs	2.69
	109550	FD1534	Hs.26981	ESTs	1.19
60	109613	FD3031	Hs.27519	ESTs	1.01
	109837	H00656	Hs.29792	ESTs	0.81
	109893	H04768	Hs.30484	ESTs	1.44
	109984	H09594	Hs.10299	ESTs	0.62
	110099	H16568	Hs.23748	ESTs	1.01
65	110837	N30796	Hs.17424	ESTs; Weakly similar to semaphorin F [H]	1.1
	111247	N69825	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	1.26
	111341	N80935	Hs.22483	ESTs	1.57
	111510	R07856	Hs.16355	ESTs	3.96
	111737	R25410	Hs.9218	ESTs	0.97
70	113195	T57112		***yc20g11.s1 Stratagene lung (#937210)	1.22
	113238	T62979	Hs.189813	ESTs	2.27
	113540	T90496	Hs.16757	ESTs	1.06
	113552	T90889	Hs.16026	ESTs	1.16
	113506	T93093	Hs.17125	ESTs	1.48
75	113595	T96965	Hs.17948	ESTs	1.54
	113946	W84753	Hs.37896	ESTs	1.79
	114251	Z39898	Hs.21948	ESTs	1.95
	114359	Z41589	Hs.153483	ESTs; Moderately similar to H1 chloride	1.42
	115230	AA278300	Hs.182980	ESTs	2.62
80	115279	AA279760	Hs.63671	ESTs	1.79
	115566	AA398083	Hs.43977	ESTs	0.86
	115965	AA446661	Hs.173233	ESTs	0.79
	116166	AA461556	Hs.202949	KIAA1102 protein	2.29
	116279	AA486073	Hs.57362	ESTs	2.27
	117023	H88157	Hs.41105	ESTs	1.36

	117209	H99959	Hs.42768	ESTs	1.46	0.48
	118901	N90719	Hs.94445	ESTs	1.51	1
	118981	N93839	Hs.39288	ESTs	1.34	0.48
5	119073	R32894	Hs.45514	v-ets avian erythroblastosis virus E26 o	1.14	0.27
	119221	R98105		**yr30g11.s1 Soares fetal liver spleen	1.32	0.53
	119824	W74536	Hs.184	advanced glycosylation end product-speci	1	0.19
	119861	W80715		ESTs; Moderately similar to IIII ALU SUB	1.83	0.45
	120041	W82775	Hs.59368	ESTs	1.23	0.55
10	120132	Z38839	Hs.125019	ESTs; Highly similar to KIAA0886 protein	0.91	0.37
	120457	AA251579	Hs.187628	ESTs	1.87	1.91
	121314	AA402799	Hs.182538	ESTs	1.3	0.31
	121643	AA417078	Hs.193767	ESTs	2.31	0.68
	121690	AA418074	Hs.110286	ESTs	1.47	0.51
15	122633	AA454080	Hs.34853	inhibitor of DNA binding 4; dominant neg	1.31	0.63
	123978	C20653	Hs.170278	ESTs	1.52	0.32
	124214	H58608	Hs.151323	ESTs	0.93	0.35
	124357	N22401		**yw37g07.s1 Morton Fetal Cochlea Homo	1.29	1
	124438	N40188	Hs.102550	ESTs	1.36	0.7
20	125167	W45560	Hs.102541	ESTs	1.45	0.69
	125174	W51835	Hs.231082	EST	3.07	3.76
	125422	AA903229	Hs.153717	ESTs	1.34	0.3
	125561	AA17667	Hs.22978	ESTs	1.89	0.63
	125631	D60988		***HUM145B09B Clontech human fetal brain	0.94	0.36
25	127002	R35380	Hs.24979	ESTs	3.02	4.06
	127307	AA369367	Hs.126712	ESTs; Weakly similar to pIL2 hypothetica	1.01	0.69
	127609	AA622559	Hs.150318	ESTs	1.21	0.32
	127959	AJ302471	Hs.124292	ESTs	2.5	1
	128458	D52193	Hs.56340	ESTs	1.13	0.33
30	128624	AA479209	Hs.102647	ESTs	1.45	0.58
	128789	AA486567	Hs.105695	ESTs	1.1	0.34
	128798	AF014958	Hs.105938	chemokine (C-C motif) receptor-like 2	1.16	0.55
	128952	R51076	Hs.107361	ESTs; Highly similar to Rap2 interacting	2.04	2.4
	129057	X62466	Hs.214742	CDW52 antigen (CAMPATH-1 antigen)	1.77	0.73
35	129210	AA401654	Hs.202949	KIAA1102 protein	1.11	0.36
	129240	W24360	Hs.237868	Interleukin 7 receptor	0.91	0.41
	129402	T63781		**yc21g01.s1 Stratagene lung (#937210)	1.36	0.43
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1	0.67	0.08
40	129583	AA487015	Hs.98314	Homo sapiens mRNA; cDNA DKFZp566L0120 (I	1.3	0.42
	129626	AA447410	Hs.11712	ESTs; Weakly similar to IIII ALU SUBFAMI	1.28	0.46
	129699	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ubliq	1.58	1
	129898	N48595	Hs.13256	ESTs	1.13	0.53
	129958	L20591	Hs.1378	annexin A3	0.81	0.31
45	130273	U59914	Hs.153863	MAD (mothers against decapentaplegic; Dr	0.59	0.22
	130655	N92934	Hs.17409	cysteine-rich protein 1 (intestinal)	1.44	0.76
	130657	T94452	Hs.201591	ESTs	0.96	0.42
	131061	N64328	Hs.22567	ESTs; Moderately similar to HYPOTHETICAL	1.51	0.45
	131066	F09006	Hs.22588	ESTs	0.97	0.37
	131263	R38334	Hs.24950	regulator of G-protein signalling 5	2.34	2.82
50	131589	U52100	Hs.29191	epithelial membrane protein 2	1.2	0.62
	131686	AA157428	Hs.30687	Grb2-associated binder 2	0.95	0.38
	131751	H18335	Hs.31562	ESTs	1.47	0.52
	132430	T23630	Hs.258675	EST	1.86	2.09
	132476	N67192	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	1.73	0.58
55	132836	F09557	Hs.57929	still (Drosophila) homolog 3	0.91	0.29
	133120	X64559	Hs.65424	telranectin (plasminogen-binding protein	0.82	0.2
	133488	D45370	Hs.74120	adipose specific 2	1.29	0.48
	133565	H57056	Hs.204831	ESTs	2.25	0.57
	133651	U97105	Hs.173381	dihydropyrimidinase-like 2	1.65	0.62
60	133835	AA059489	Hs.76640	ESTs; Highly similar to RGC-32 [R.norveg	1.16	0.34
	133978	W73859	Hs.78061	transcription factor 21	0.79	0.27
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec	0.99	0.28
	134299	AA487558	Hs.8135	ESTs	1.02	0.46
	134300	U81984	Hs.166082	endothelial PAS domain protein 1	0.86	0.42
65	134323	AA028976	Hs.8175	Homo sapiens mRNA; cDNA DKFZp564M0763 (I	1.19	0.27
	134343	O50683	Hs.82028	transforming growth factor; beta recepto	1.21	0.67
	134417	D87969	Hs.82921	solute carrier family 35 (CMP-sialic aci	1.28	1
	134561	U76421	Hs.85302	adenosine deaminase; RNA-specific; B1 (h	2.12	0.55
	134624	W67147	Hs.8700	deleted in liver cancer 1	2.35	2.74
70	134696	H88354	Hs.8861	ESTs	1.35	0.33
	134749	L10955	Hs.89485	carbonic anhydrase IV	0.89	0.2
	134786	L06139	Hs.89640	TEK tyrosine kinase; endothelial (venous	0.48	0.21
	134869	T35288	Hs.90421	ESTs; Moderately similar to IIII ALU SUB	2.14	2.64
	135346	M21056	Hs.992	phospholipase A2; group IB (pancreas)	0.63	0.13
75	100113	D00591	Hs.84746	Chromosome condensation 1	1	2.15
	100147	D13666	Hs.136348	Homo sapiens mRNA for osteoblast specifi	0.5	2
	100280	D42085	Hs.155314	KIAA0095 gene product	1.02	1.39
	100335	D63391	Hs.6793	platelet-activating factor acetylhydrola	1	5.58
	100360	D78335	Hs.75939	Uridine monophosphate kinase	0.91	2.04
80	100372	D79997	Hs.184339	KIAA0175 gene product	0.75	2.03
	100486	HG1112-HT1112		TIGR: ras-like protein TC4	1.09	1.93
	100559	HG2197-HT2267		*collagen, type VII, alpha 1*	0.97	3.6
	100576	HG2290-HT2386		*calitonin/alpha-CGRP, alt. transcript	1	1
	100668	HG2981-HT3938		*TIGR: CD44 (epican, alt. transcript 12	0.85	1.9
85	100906	HG4716-HT5158		Guanosine 5'-Monophosphate Synthase	1.18	2.29
	100930	HG721-HT4827		*TIGR: placental protein 14, endometrial	1	1.45

	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	0.84	2.6
	101031	J05070	Hs.151738	*Matrix metalloproteinase 9 (gelatinase	0.77	1.52
	101111	L08424	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
5	101124	L10343	Hs.112341	*Protease inhibitor 3, skin-derived (SKA	0.62	2.67
	101175	L18920	Hs.36980	*Melanoma antigen, family A, 2'	1	1
	101204	L24203	Hs.82237	Ataxia-telangiectasia group D-associated	0.74	4.1
	101431	M19888	Hs.1076	Small proline-rich protein 1B (comifin)	0.85	2.51
	101448	M21389	Hs.195850	keratin 5 (epidermolysis bullosa simplex	0.61	8.83
10	101511	M27826	Hs.267319	Endogenous retroviral protease	1.03	1.13
	101526	M29540	Hs.220529	Carcinoembryonic antigen-related cell ad	1.07	4.61
	101548	M31328	Hs.71642	*Guanine nucleotide binding protein (G p	0.97	1.13
	101625	M57293		*Human parathyroid hormone-related pepti	1	1
	101649	M50047	Hs.1690	Heparin-binding growth factor binding pr	1	2.7
	101724	M69225	Hs.620	bullosa pemphigoid antigen 1 (230/240kD)	1	8.98
15	101748	M76482	Hs.1925	Desmoglein 3 (pemphigus vulgaris antigen	1	2.78
	101759	M80244	Hs.184601	*Solute carrier family 7 (cationic amino	1.07	2.45
	101804	M86699	Hs.169840	TTK protein kinase	1	1
	101806	M86757	Hs.112408	S100 calcium-binding protein A7 (psorias	0.74	1.76
	101809	M86849		*Homo sapiens connexin 26 (GJB2) mRNA, c	1	7
20	101845	M93426	Hs.78867	*Protein tyrosine phosphatase, receptor-	1	1
	101851	M94250	Hs.82045	Midkine (neurite growth-promoting factor	1.13	2.6
	102083	U10323	Hs.75117	*Interleukin enhancer binding factor 2,	1.03	1.61
	102154	U17760	Hs.75517	*Laminin, beta 3 (nicein (125kD), kalin)	0.94	3.62
	102193	U20758	Hs.313	secreted phosphoprotein 1 (osteopontin;	0.34	4.59
25	102305	U33286	Hs.90073	chromosome segregation 1 (yeast homolog)	1.45	2.97
	102348	U37519	Hs.87539	Aldehyde dehydrogenase 8	0.52	2.25
	102581	U61145	Hs.77256	Enhancer of zeste (Drosophila) homolog 2	0.91	2.46
	102610	U65011	Hs.30743	Preferentially expressed antigen in mela	1	3.88
	102623	U66083	Hs.37110	*Melanoma antigen, family A, 9 (MAGE-9)	1	1
30	102669	U71207	Hs.29279	Eyes absent (Drosophila) homolog 2	1	1
	102696	U74612	Hs.239	Forkhead box M1	1.06	2.77
	102829	U91618	Hs.80962	Neurotensin	1	1
	102888	X04741	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.13	2.59
	102913	X07696	Hs.80342	keratin 15	0.7	4.72
35	102915	X07820	Hs.2258	Matrix Metalloproteinase 10 (Stromolysin	1.15	3.35
	102963	X15943	Hs.37058	*Calcitonin/calcitonin-related polypepti	1	1
	103021	X53587	Hs.85266	*Integrin, beta 4*	1.38	2.34
	103036	X54925	Hs.83169	Matrix metalloprotease 1 (interstitial c	1	14.93
40	103058	X57348	Hs.184510	Stratfin	1.25	4.17
	103060	X57766	Hs.155324	matrix metalloproteinase 11 (stromelysin	1	1.72
	103119	X63629	Hs.2877	*Cadherin 3, P-cadherin (placental)	1.16	7.38
	103206	X72755	Hs.77367	monokine induced by gamma interferon	0.71	1.48
	103242	X76342	Hs.389	*Alcohol dehydrogenase 7 (class IV), mu	1	1
45	103312	X82693	Hs.3185	*Lymphocyte antigen 6 complex, locus D;	0.92	1.28
	103478	Y07755	Hs.38991	S100 calcium-binding protein A2	1.05	5.81
	103558	Z19574	Hs.2785	keratin 17	0.65	6.68
	103576	Z26317	Hs.2631	Desmoglein 2	0.79	1.73
	103587	Z29083	Hs.82128	ST4 Oncofetal antigen	1	3.93
50	103594	Z31560	Hs.816	*SRY (sex determining region Y)-box 2, p	0.71	7.23
	103768	AA089997		*ESTs, Highly similar to integral membra	0.99	1.8
	104158	AA454908	Hs.8127	KIAA0144 gene product	0.96	1.29
	104558	R56578	Hs.88959	Human DNA sequence from clone 957N21 on	1.23	7.23
	104689	AA010665		ESTs	0.96	2.11
55	104733	AA019498	Hs.23071	ESTs	1.18	1.88
	104906	AA055809	Hs.26802	Protein kinase domains containing protei	1.11	3.15
	104978	AA088458	Hs.19322	ESTs; Weakly similar to H11 ALU SUBFAM1	1.64	2.89
	105012	AA116036	Hs.9329	*Homo sapiens mRNA for fls353, complete	1.19	3.91
	105175	AA186804	Hs.25740	ESTs; Weakly similar to unknown [S.cerev	0.9	4.63
60	105263	AA227926	Hs.6682	ESTs	0.95	2.87
	105298	AA233459	Hs.26369	ESTs	1	1.13
	105312	AA233854	Hs.23348	S-phase kinase-associated protein 2 (p45	1.32	3.01
	105719	AA291644	Hs.36793	Hypothetical protein FLJ23188	1.28	2.31
	105743	AA293300	Hs.9598	ESTs	1	1
65	106012	AA411621	Hs.8895	ESTs; same as BFH67	0.94	2.04
	106231	AA429571	Hs.38002	KIAA1355 protein	1.04	1.5
	106540	AA454607	Hs.38114	Hypothetical protein FLJ11100	1.26	2.26
	106575	AA456039	Hs.105421	ESTs	1	2
	106632	AA459897	Hs.11950	GPI-anchored metastasis-associated prote	0.87	1.32
70	106727	AA465342	Hs.34045	Hypothetical protein FLJ20764	0.87	1.59
	106908	AA490237	Hs.222024	Transcription factor BMAL2 (cycle-like f	0.61	1.6
	107059	AA608545	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	0.48	2.67
	107104	AA609786	Hs.15243	Nucleolar protein 1 (120kD)	1.01	1.44
	107151	AA621169	Hs.8687	ESTs; procollagen I-N proteinase	0.97	2.89
75	107284	S74039	Hs.291904	Accessory proteins BAP31/BAP29	1.15	3.65
	107901	AA026418	Hs.91539	ESTs	0.72	3.44
	107922	AA028028	Hs.61460	Ig superfamily receptor LNIR precursor	1	2.48
	107932	AA029317	Hs.18878	Hypothetical protein FLJ21620	1	1
	108695	AA121315	Hs.70823	KIAA1077 protein	0.91	3.53
80	108857	AA133250	Hs.62180	ESTs	1	1
	108860	AA133334	Hs.129911	ESTs	0.73	7.3
	108990	AA152296	Hs.72045	ESTs	1	1
	109166	AA179845	Hs.73625	*RAB6 interacting, kinesin-like (rabkine	1	4.55
	109424	AA227919	Hs.85962	Hyaluronan synthase 3	1	1.28
85	109665	F05012	Hs.27027	Hypothetical protein DKFZp762H1311	1.42	2
	109970	H09281	Hs.13234	ESTs	1.13	2.16

5	110015	H10998	Hs.7164	A disintegrin and metalloproteinase doma	0.84	1.95
	110156	H18957	Hs.4213	ESTs	0.94	1.41
	110561	H59617	Hs.5199	HSPC150 protein similar to ubiquitin-con	0.91	3.18
	111223	N68921	Hs.34806	ESTs; Weakly similar to neogenin [H.sapi	0.91	3.13
	111345	N89820	Hs.14559	Hypothetical protein FLJ10540	1	1.25
10	111876	R38239	Hs.293246	*ESTs, Weakly similar to putative p150 [0.83	1.27
	111902	R39191	Hs.109445	KIAA1020 protein	0.91	0.91
	112244	R51309	Hs.70823	KIAA1077 protein	0.77	3.01
	112973	T17271		*cDNA FLJ13308 fis, clone OVARC1001436,	1	1
	112989	T23482	Hs.89981	"Diacylglycerol kinase, zeta (104kD)"	0.55	1.03
15	113047	T25867	Hs.7549	ESTs	0.87	2
	113095	T40920	Hs.126733	ESTs	1	1
	113531	T90345	Hs.16740	Hypothetical protein FLJ11035	0.42	1.44
	113970	W86748	Hs.8109	ESTs	1.17	1.73
	114346	Z41450	Hs.130489	*ATPase, aminophospholipid transporter-I	0.86	0.82
20	114407	AA010188	Hs.103305	ESTs	0.8	1.88
	114471	AA028074	Hs.104613	RP42 homolog	1.06	1.34
	114509	AA043551	Hs.101799	KIAA1350 protein	1.82	2.32
	115060	AA253214	Hs.198249	*Gap junction protein, beta 5 (connexin	0.79	1.49
	115091	AA255900	Hs.184523	KIAA0955 protein	0.72	1.92
25	115123	AA256642	Hs.236894	*ESTs, High sim to LRP1_hu low density I	0.59	1.97
	115291	AA279943	Hs.122579	ESTs	1	1.25
	115506	AA292537	Hs.45207	Hypothetical protein KIAA1335	1.15	1.48
	115522	AA331393	Hs.47378	ESTs	0.5	3.29
	115536	AA347193	Hs.62180	ESTs	1	1
30	115697	AA411502	Hs.63325	Homo sapiens type II membrane serine pro	1	6.53
	115909	AA436666	Hs.59761	ESTs	1	6.98
	115978	AA447522	Hs.69517	Differentially expressed in Fanconi anem	1	2.31
	116028	AA452112	Hs.42644	thioredoxin-like	0.99	1.68
	116107	AA456958	Hs.92030	ESTs	1.14	1.8
35	116134	AA460246	Hs.50441	CGI-04 protein	1.11	1.86
	116157	AA461063	Hs.44298	Hypothetical protein	0.99	1.9
	116158	AA461187	Hs.61762	Hypoxia-inducible protein 2	0.44	0.86
	116335	AA495830	Hs.87013	*Homo sapiens cDNA FLJ10238 fis, clone H	0.62	3.89
	116483	C14092	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.04	2.36
40	117320	N23239	Hs.211092	LUNX protein; PLUNC(palate lung & nasal	0.51	0.64
	117557	N33920	Hs.44532	Diubiquitin	1.11	2.63
	117693	N40939	Hs.112110	PTD007 protein	0.98	1.79
	117881	N50073	Hs.260622	Butyrate-induced transcript 1	1	1.43
	118368	N64339	Hs.48956	ESTs	0.67	2.86
45	118566	N68558	Hs.42824	Hypothetical protein FLJ10718	1.21	0.83
	118695	N71781	Hs.50081	KIAA1199 see CVA7.doc	0.88	1.63
	119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical pro	1	1
	119845	W79920	Hs.58561	G protein-coupled receptor 87	1	1
	120102	W95428	Hs.132927	*ESTs, Moderately similar to p53 regulat	1	1
50	120104	W95477	Hs.180479	ESTs	0.69	3.07
	120486	AA253400	Hs.137569	Tumor protein 63 kDa with strong homolog	1.08	12.05
	120859	AA350158	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
	120880	AA360240	Hs.97019	EST	1	1
	120948	AA397822	Hs.104650	Hypothetical protein FLJ10292	1.04	2.15
55	120983	AA398209	Hs.97587	EST	1	1
	121362	AA405500	Hs.97932	Chondromodulin I precursor	1	1
	121369	AA405657	Hs.128791	CGI-09 protein	1	1.8
	121791	AA423978	Hs.293317	*ESTs, Weakly similar to JM27 [H.sapiens	1	1
	123005	AA479726	Hs.105577	ESTs	1	1
60	123044	AA481549	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	0.95	1.88
	123160	AA488687	Hs.284235	ESTs	1.59	4.98
	123479	AA599469	Hs.135056	clone RP5-850E9 on chromosome 20	1.19	1.64
	123571	AA608956	Hs.112619	*ESTs, Weakly similar to PQ0109 Purkinje	1.03	1.14
	123829	AA620697	Hs.112208	XAGE-1 protein	1.39	2.2
65	124006	D60302	Hs.108977	ESTs	1	4.85
	124059	F13673	Hs.99769	ESTs	1.49	8.62
	124950	T15386	Hs.194766	Seizure related gene 6 (mouse)-like	0.76	0.77
	125218	W73561	Hs.110024	NADH:ubiquinone oxidoreductase MLRQ subu	1.33	1.77
	125453	R06041	Hs.18048	*Melanoma antigen, family A, 10"	0.8	1.42
70	125759	AA425587	Hs.82226	Glycoprotein (transmembrane) numb	1.52	2.26
	125972	AA434562	Hs.35406	*ESTs, Highly similar to unnamed protein	1.05	2.48
	125994	H55782	Hs.270799	EST	1	1.95
	126395	N70192	Hs.278956	Hypothetical protein FLJ12929	1	1.35
	126645	A1157942	Hs.61635	STEAP1 (Homo sapiens BAC clone RG041D11	1	2.23
75	127221	A1354332	Hs.72365	ESTs	0.73	3.27
	127479	AA513722	Hs.179729	collagen; type X; alpha 1 (Schmid metaph	0.51	1.94
	128192	A1204246		KIAA1085 protein	1.8	3.16
	128610	L38608	Hs.10247	activated leucocyte cell adhesion molecu	0.89	0.97
	128777	U46006	Hs.10526	Cysteine and glycine-rich protein 2	1	1
80	128924	AA234962	Hs.26557	Plakophilin 3	1.3	2.97
	129041	H58873	Hs.169902	*Solute carrier family 2 (facilitated gl	0.84	2.04
	129099	H50398	Hs.108660	*ATP-binding cassette, sub-family C (CFT	0.87	1.04
	129404	AA172056	Hs.111128	ESTs	1	1
	129466	L42583		*Genbank Homo sapiens keratin 6 isoform	0.72	12.67
85	129605	S72493	Hs.115947	Keratin 16 (focal non-epidermolytic palm	0.92	1.5
	129628	U26727	Hs.1174	*Cyclin-dependent kinase inhibitor 2A (m	0.85	1.93
	130023	X13461	Hs.239600	Calmodulin-like 3	0.84	1.22
	130080	X14850	Hs.147097	*H2A histone family, member X"	0.98	1.96
	130385	AA126474	Hs.155223	stanniocalcin 2	1	1

	130410	V01514	Hs.155421	Alpha-fetoprotein	0.63	0.63
	130441	U35835	Hs.301387	"Human DNA-PK mRNA, partial cds"	1.15	3.65
	130482	L32866	Hs.1578	Baculoviral IAP repeat-containing 5 (sur	1	1.88
5	130553	AA430032	Hs.252587	Pituitary tumor-transforming 1	0.92	1.96
	130577	M35410	Hs.162	Insulin-like growth factor binding prote	1.17	4.7
	130627	L23808	Hs.1695	Matrix metalloproteinase 12 (macrophage	0.69	4.05
	130800	AA223386	Hs.19574	ESTs; Weakly similar to katanin p80 subu	1.13	2.41
	130839	AA598689	Hs.21400	ESTs	0.8	0.89
10	131046	X02530	Hs.2248	INTERFERON-GAMMA INDUCED PROTEIN PRECURS	0.8	1.15
	131244	D38076	Hs.24763	RAN binding protein 1	1.13	1.85
	131877	J04088	Hs.156346	Topoisomerase (DNA) II alpha (170kD)	1	1
	131927	AA481549	Hs.34780	"Doublecortin; lissencephaly, X-linked (0.81	0.62
	131965	W90146	Hs.35962	ESTs	0.74	3.27
15	131978	D80008	Hs.36232	KIAA0186 gene product	1	1
	132354	L05187	Hs.211913	Small proline-rich protein 1A	0.69	1.43
	132543	AA417152	Hs.5101	ESTs; Highly similar to protein regulati	0.79	4.27
	132632	N59764	Hs.5398	guanine-monophosphate synthetase	1	1.08
	132653	U31201	Hs.54451	"laminin gamma2 chain gene (LAMC2), exon	1	1
20	132659	Z75190	Hs.54481	"Low density lipoprotein receptor-relate	0.89	0.89
	132710	W93726	Hs.55279	"Serine (or cysteine) proteinase inhibit	0.64	4.41
	132758	W52432	Hs.56105	"ESTs, Weakly similar to WDNM RAT WDNM1	1.55	2.08
	132767	L05188	Hs.231622	Small proline-rich protein 2B	0.83	1.66
	132816	M74542	Hs.575	Aldehyde dehydrogenase 3	0.55	0.55
25	132990	AA458761	Hs.18387	transcription factor AP-2 alpha (activat	1	3.53
	133070	U69611	Hs.64311	"A disintegrin and metalloproteinase dom	1.16	2
	133282	U52960	Hs.286145	"SRB7 (suppressor of RNA polymerase B, y	1	2.7
	133317	AA215299	Hs.70830	U6 snRNA-associated Sm-like protein Lsm7	0.95	1.42
	133370	AA156897	Hs.72157	Homo sapiens mRNA; cDNA DKFZp56411922	1.12	2.55
30	133391	X57579	Hs.727	H.sapiens activin beta-A subunit (exon 2	1.65	1.76
	133832	H03387	Hs.241305	estrogen-responsive B box protein (EBBP)	1.02	1.39
	134032	Z81326	Hs.78589	"Serine (or cysteine) proteinase inhibit	1	1
	134168	AA398908	Hs.181634	"Homo sapiens cDNA: FLJ23502 tis, clone	0.95	1.53
	134218	AA227480	Hs.80205	Pim-2 oncogene	1.36	2.48
35	134405	R67275	Hs.82772	"collagen, type XI, alpha 1""	0.76	2.86
	134453	X70683	Hs.83484	SRV (sex determining region Y)-box 4	1.89	3.78
	134470	X54942	Hs.83758	CDC28 protein kinase 2	1.82	4.11
	134645	U87459	Hs.167379	"Cancer/testis antigen (NY-ESO-1, CTAG1,	0.82	0.83
	134781	M17183	Hs.89626	Parathyroid hormone-like hormone	1	1
40	135002	U19147	Hs.272484	G antigen 6	1	1
	100040	M97935		AFFX control: STAT1	0.92	1.25
	101201	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin;	2.92	8.5
	101664	M60752	Hs.121017	H2A histone family; member A	1	1
	102025	U03911	Hs.78934	mutS (E. coli) homolog 2 (colon cancer;	0.8	1.61
45	102031	U04898	Hs.2156	RAR-related orphan receptor A	1	1
	102221	U24576		LIM domain only 4	1	1
	102270	U30255	Hs.75888	phosphogluconate dehydrogenase	1.08	1.43
	102339	U37022	Hs.95577	cyclin-dependent kinase 4	0.88	1.32
	102391	U41668	Hs.77494	deoxyguanosine kinase	1.07	1.58
50	103000	X51956	Hs.146580	enolase 2; (gamma; neuronal)	0.91	1.49
	103395	X94754	Hs.119503	methionine-tRNA synthetase	0.89	1.32
	105638	AA281599	Hs.20418	Homo sapiens mRNA for for histone H2B; c	0.91	1.25
	105726	AA292328	Hs.9754	activating transcription factor 5	0.94	1.48
	114841	AA234722	Hs.55408	ESTs; Moderately similar to CALCIUM-DEPE	0.78	1.56
55	115206	AA262491	Hs.186572	ESTs	1	1
	115906	AA436616	Hs.82302	ESTs	0.74	2.52
	119132	R49046	Hs.107911	ATP-binding cassette; sub-family B (MDR/	1.1	1.51
	124163	H30539	Hs.189838	ESTs	1	1
	126487	AA482505	Hs.184601	solute carrier family 7 (cationic amino	1.01	1.46
60	127141	AA307980	Hs.75478	KIAA0956 protein	0.85	1.4
	128034	AA905754	Hs.75103	tyrosine 3-monooxygenase/tryptophan 5-mo	1	1.18
	128609	AA234365	Hs.102456	survival of motor neuron protein Interac	1	1.5
	128895	R37753	Hs.106885	ESTs	1.7	2
	130199	Z48579	Hs.172028	a disintegrin and metalloprotease domain	1	1
65	130524	U89995	Hs.159234	forkhead box E1	1	1
	133000	U24152	Hs.82402	p21/Cdc42/Rac1-activated kinase 1 (yeast	1	1
	133658	M25756	Hs.75426	secretogranin II (chromogranin C)	1	1
	135047	AA460466	Hs.93597	ESTs	1	1
	100053	M27830		AFFX control: 28S ribosomal RNA	0.88	1.53
70	100114	D00596	Hs.82962	thymidylate synthetase	0.68	1.86
	100128	D11094	Hs.61153	proteasome (prosome; macropain) 26S subu	1.29	2.03
	100154	D14657	Hs.81892	KIAA0101 gene product	0.71	4.26
	100161	D14694	Hs.77329	phosphatidylserine synthase 1	1.02	1.56
	100168	D14874	Hs.394	adrenomedullin	0.46	1.17
75	100187	D17793	Hs.78183	aldo-keto reductase family 1; member C3	1	1
	100188	D21063	Hs.57101	minichromosome maintenance deficient (S.	0.97	1.4
	100217	D26600	Hs.89545	proteasome (prosome; macropain) subunit;	1.13	1.9
	100220	D28364		"Human mRNA for annexin II, 5'UTR (seq	1.11	1.53
	100287	D43950	Hs.1600	chaperonin containing TCP1; subunit 5 (e	1.13	2.09
80	100297	D49489	Hs.182429	protein disulfide isomerase-related prot	0.92	1.78
	100330	D55716	Hs.77152	minichromosome maintenance deficient (S.	1.07	1.61
	100355	D78129		"Homo sapiens mRNA for squalene epoxid	0.96	1.87
	100364	D78586	Hs.154868	carbamoyl-phosphate synthetase 2; aspart	1.49	2.46
	100368	D79987	Hs.153479	extra spindle poles; S. cerevisiae; homo	0.59	1.32
85	100398	D84557	Hs.155462	minichromosome maintenance deficient (mi	1.08	1.9
	100438	D87448	Hs.91417	topoisomerase (DNA) II binding protein	1	2.15

	100455	D87953	Hs.75789	N-myc downstream regulated	0.91	1.48
	100491	HG1153-HT1153		Nucleoside Diphosphate Kinase Nm23-H2s	0.99	1.41
	100518	HG174-HT174		Desmoplakin I	1.28	3.17
5	100528	HG1828-HT1857		***Nexin, Glia-Derived***	0.68	1.9
	100661	HG2874-HT3018		Ribosomal Protein L39 Homolog	1.1	5.44
	100667	HG2981-HT3127		***Epitan, Alt. Splice 11***	0.8	1.97
	100830	HG4074-HT4344		Rad2	1.01	2.12
	101061	K03515	Hs.944	glucose phosphate isomerase	0.91	1.79
10	101131	L10838	Hs.167460	splicing factor; arginine/serine-rich 3	1.23	1.87
	101162	L14595	Hs.174203	solute carrier family 1 (glutamate/neutr	1.35	2.73
	101181	L19686	Hs.73798	macrophage migration inhibitory factor (1.03	1.78
	101183	L19779	Hs.795	H2A histone family; member O	0.57	1.3
	101216	L25876	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	0.7	2.2
15	101228	L27705	Hs.82916	chaperonin containing TCP1; subunit 6A (0.99	1.99
	101233	L29008	Hs.878	sorbitol dehydrogenase	0.82	2.11
	101247	L33801	Hs.78802	glycogen synthase kinase 3 beta	1.2	1.91
	101332	L47276		***Homo sapiens (cell line HL-6) alpha t	0.69	2.78
	101342	L76191	Hs.182018	interleukin-1 receptor-associated kinase	1.04	1.84
	101396	M15796	Hs.78996	proliferating cell nuclear antigen	0.95	3.55
20	101423	M18391	Hs.89839	EphA1	1	1.5
	101445	M21259	Hs.1066	small nuclear ribonucleoprotein polypept	1.21	1.96
	101505	M27396	Hs.75692	asparagine synthetase	0.93	1.6
	101525	M29536	Hs.12163	eukaryotic translation initiation factor	1.19	1.93
25	101535	M30448	Hs.251669	casein kinase 2; beta polypeptide	0.96	1.42
	101607	M38690	Hs.1244	CD9 antigen (p24)	1.11	1.25
	101624	M56998		***Human alpha-1 collagen type I gene, 3	1.17	1.98
	101758	M77836	Hs.79217	pyroline-5-carboxylate reductase 1	1.77	3.45
	101839	M93036	Hs.692	membrane component; chromosomal 4; surfa	0.71	1.45
	101853	M94362	Hs.76084	lamin B2	0.84	1.19
30	101977	S83364		***putative Rab5-interacting protein (cl	0.89	1.9
	101992	U01038	Hs.77597	polo (Drosophila)-like kinase	0.66	1.46
	102009	U02680	Hs.82643	protein tyrosine kinase 9	1.23	3.35
	102012	U03057	Hs.118400	singed (Drosophila)-like (sea urchin fas	0.85	1.88
35	102039	U05861	Hs.201967	aldo-keto reductase family 1; member C1	0.93	2.32
	102123	U14518	Hs.1594	centromere protein A (17kD)	1	4.28
	102130	U15009	Hs.1575	small nuclear ribonucleoprotein D3 polyp	0.89	1.42
	102148	U16954	Hs.75823	ALL-1-fused gene from chromosome 1q	0.8	2.95
	102210	U23028	Hs.2437	eukaryotic translation initiation factor	1.01	1.34
40	102220	U24389	Hs.65436	lysyl oxidase-like 1	1.15	2.34
	102260	U28386	Hs.159557	karyopherin alpha 2 (RAG cohort 1; impor	1.14	2.69
	102330	U35451	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta	1.05	1.7
	102423	U44754	Hs.179312	small nuclear RNA activating complex; po	1.14	2.99
	102455	U48705	Hs.75562	discoidin domain receptor family; member	1.05	2.01
45	102499	U51478	Hs.76941	ATPase; Na+/K+ transporting; beta 3 poly	1.27	1.92
	102522	U53347	Hs.183556	solute carrier family 1 (neutral amino a	0.84	1.31
	102590	U62136		***Homo sapiens enterocyte differentiati	1.11	1.6
	102676	U72514	Hs.12045	putative protein	1.04	2.17
	102687	U73379	Hs.93002	ubiquitin carrier protein E2-C	0.86	2.28
50	102704	U76638	Hs.54089	BRCA1 associated RING domain 1	1.12	1.63
	102781	U83843		***Human HIV-1 Nef interacting protein (0.9	1.39
	102784	U85658	Hs.61796	transcription factor AP-2 gamma (activat	0.98	2.16
	102827	U91327	Hs.6456	chaperonin containing TCP1; subunit 2 (b	0.96	1.62
	102935	X13482	Hs.80506	small nuclear ribonucleoprotein polypept	1.21	4.2
55	102972	X16662	Hs.87268	annexin A8	1.25	2.32
	102983	X17620	Hs.118638	non-metastatic cells 1; protein (NM23A)	1.03	1.83
	103023	X53793	Hs.117950	multifunctional polypeptide similar to S	1.58	5.44
	103038	X54941	Hs.77550	CDC28 protein kinase 1	1.32	3.79
	103075	X59543	Hs.2934	ribonucleotide reductase M1 polypeptide	1.11	2.58
60	103168	X68314	Hs.2704	glutathione peroxidase 2 (gastrointestin	0.75	3.05
	103185	X89910	Hs.74368	transmembrane protein (63kD); endoplasm	1.01	1.97
	103212	X73874	Hs.2393	phosphorylase kinase; alpha 1 (muscle)	0.95	1.72
	103223	X74801	Hs.1708	chaperonin containing TCP1; subunit 3 (g	0.97	1.77
	103260	X78416	Hs.3155	casein; alpha	1	1
65	103262	X78565	Hs.204133	hexabrachion (tenascin C; cytotoxic)	1.23	3.09
	103330	X85373	Hs.77496	small nuclear ribonucleoprotein polypept	1.12	2.25
	103364	X90872	Hs.75854	SULT1C sulfotransferase	2.85	4.62
	103375	X91868	Hs.54416	sine oculis homeobox (Drosophila) homolo	1	2.48
	103391	X94453	Hs.114366	pyroline-5-carboxylate synthetase (glut	1	1.53
70	103404	X95586	Hs.78596	proteasome (prosome; macropain) subunit;	0.92	1.53
	103437	X98260	Hs.82254	M-phase phosphoprotein 11	0.92	1.54
	103448	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	0.55	0.96
	103605	Z35402	Hs.194657	cadherin 1; E-cadherin (epithelial)	1.32	2.51
	103646	Z68228	Hs.2340	junction plakoglobin	0.88	1.28
75	103658	Z74615	Hs.172928	collagen; type I; alpha 1	1.06	2.98
	103774	AA092898	Hs.92918	ESTs; Weakly similar to R07G3.8 [C.elega	1.88	4.66
	104261	AF008442	Hs.5409	RNA polymerase I subunit	0.87	2.17
	104276	C02193	Hs.85222	ESTs; Weakly similar to R27090_2 [H.sapi	1.4	2.49
	104289	C16281	Hs.75478	KIAA0956 protein	1.15	1.68
	104434	L02870	Hs.1640	collagen; type VII; alpha 1 (epidermolys	1.04	1.49
80	104453	M19169	Hs.123114	cystatin SN	0.38	0.76
	104611	R98280	Hs.125845	ribulose-5-phosphate-3-epimerase	1.08	2.25
	104758	AA024661	Hs.7010	ESTs; Weakly similar to ACYL-COA DEHYDRO	1.14	1.65
	105114	AA156532	Hs.11801	adenosine A2b receptor pseudogene	0.91	1.38
85	105132	AA159501	Hs.247280	HBV associated factor	1.08	1.7
	105174	AA186613	Hs.34744	ESTs	0.95	2.05

	105280	AA232215	Hs.14600	ESTs	1	1.4
	105344	AA235303	Hs.8645	ESTs	0.72	2.02
	105516	AA257971	Hs.21214	ESTs	1.35	3.56
5	105621	AA280865	Hs.6375	Homo sapiens mRNA; cDNA DKFZp564K0222 (f	1.23	1.82
	105698	AA287393	Hs.15202	ESTs; Weakly similar to oligodendrocyte-	0.98	1.28
	105705	AA290767	Hs.101282	Homo sapiens mRNA; cDNA DKFZp434B102 (fr	0.92	1.32
	105724	AA292098	Hs.22934	ESTs; Weakly similar to ZINC FINGER PROT	0.99	1.41
	105782	AA350215	Hs.21580	ESTs	1	1
	105799	AA372018	Hs.24743	ESTs	1.08	1.78
10	105807	AA393803	Hs.16869	ESTs; Moderately similar to COLLAGEN ALP	0.95	1.34
	105891	AA400768	Hs.26562	ESTs; Weakly similar to tumor necrosis f	0.87	2.25
	105936	AA404338		ESTs	1.14	1.46
	106069	AA417741	Hs.29899	ESTs; Weakly similar to ZINC FINGER PROT	1	1
	106103	AA421104	Hs.12094	ESTs	1.04	1.44
15	106140	AA424524	Hs.14912	KIAA0286 protein	1.23	2.11
	106149	AA424881	Hs.256301	ESTs	0.83	1.48
	106154	AA425304	Hs.6994	ESTs	0.77	2.05
	106182	AA426609	Hs.10862	ESTs	0.74	2.23
20	106220	AA428582	Hs.32196	ESTs; Moderately similar to metargidin p	0.97	1.99
	106228	AA429290	Hs.17719	ESTs	0.99	1.54
	106318	AA436570	Hs.9605	pre-mRNA cleavage factor Irm (25kD)	0.95	2.09
	106341	AA441798	Hs.5243	ESTs; Moderately similar to pIL2 hypothe	0.98	2.66
	106432	AA448850	Hs.17138	ESTs	0.95	1.93
25	106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	1	1
	106483	AA451676	Hs.30299	IGF-II mRNA-binding protein 2	1.4	2.29
	106599	AA457235	Hs.12842	ESTs; Moderately similar to non-function	1	1.82
	106611	AA458904	Hs.26267	ESTs; Weakly similar to IorsinA [H.sapie	1.49	2.78
	106654	AA460449	Hs.3784	ESTs; Highly similar to phosphoserine am	1	1.4
30	107076	AA609145	Hs.21143	ESTs; Weakly similar to fos39554_1 [H.s	1.11	1.49
	107115	AA610108	Hs.27693	ESTs; Highly similar to CG1-124 protein	1	1.03
	107129	AA620553	Hs.4756	flap structure-specific endonuclease 1	1.13	3.63
	107159	AA621340	Hs.10600	ESTs; Weakly similar to ORF YKR081c [S.c	1.05	2.09
	107444	W28391	Hs.5181	proliferation-associated 2G4; 38kD	1.18	1.9
35	107481	W58247	Hs.27437	Homo sapiens kinesin superfamily motor K	0.99	2.74
	107516	X56597	Hs.99853	fibrillarin	0.94	1.77
	107529	Y12085	Hs.5092	nucleolar protein (KKE/D repeat)	1.05	2.29
	107531	Y13936	Hs.17883	protein phosphatase 1G (formerly 2C); ma	1.06	1.62
	107801	AA019433	Hs.173100	ESTs	1.03	1.4
40	107957	AA031948	Hs.57548	ESTs	0.95	1.46
	108565	AA085342	Hs.1526	ATPase; Ca++ transporting; cardiac muscl	0.59	1.35
	108780	AA128561	Hs.117938	collagen; type XVII; alpha 1	1	7.53
	108828	AA131584	Hs.71435	DKFZP564O0463 protein	1.33	2.56
	109060	AA160879	Hs.241551	chloride channel; calcium activated; fam	0.87	1.42
45	109112	AA169379	Hs.72865	ESTs	1.03	2.31
	109344	AA213696	Hs.86559	poly(A)-binding protein-like 1	0.97	1.55
	109412	AA227145	Hs.209473	ESTs; Weakly similar to REGULATOR OF MIT	0.76	1.87
	110780	N23174	Hs.22891	solute carrier family 7 (cationic amino	0.9	0.95
	110958	N50550	Hs.24587	signal transduction protein (SH3 contain	1.17	2.26
50	111018	N54067	Hs.3628	mitogen-activated protein kinase kinase	1.21	1.85
	111337	N79512	Hs.16607	ESTs; Highly similar to Myosin heavy cha	1	1.45
	112305	R54822	Hs.26244	ESTs	1	1
	112401	R61279	Hs.237536	ESTs; Weakly similar to F25B5.3 [C.elega	1.24	1.64
	112853	T02843	Hs.4351	EST	1.56	1.96
55	112869	T03313	Hs.4747	dyskeratosis congenita 1; dyskerin	1.03	1.57
	112992	T23513	Hs.7147	ESTs	1	1
	113048	T25895	Hs.184008	ESTs; Weakly similar to RNA-binding prot	1.37	2.26
	113063	T32438	Hs.5027	ESTs	1	1
	113179	T55182	Hs.152571	ESTs; Highly similar to IGF-II mRNA-bind	1.33	2.7
60	113573	T91166	Hs.16990	ESTs	0.76	1.47
	113811	W44928	Hs.4878	ESTs	0.79	1.51
	114086	Z38266	Hs.12770	Homo sapiens PAC clone DJ0777023 from 7p	0.9	1.34
	114587	AA070827	Hs.180320	ESTs; Weakly similar to GOLGI 4-TRANSMEM	1.02	1.76
	114846	AA234929	Hs.44343	ESTs	1.32	2.36
65	114964	AA243873	Hs.82184	ring finger protein 3	1.1	1.84
	115047	AA252627	Hs.22554	homeo box B5	1.01	2.36
	115166	AA258409	Hs.196907	myelin protein zero-like 1	1.05	2.31
	115167	AA258421	Hs.43728	hypothetical protein	1.52	2.52
	115239	AA278550	Hs.73291	ESTs; Weakly similar to similar to the b	0.7	2.57
70	115278	AA279757	Hs.67466	ESTs; Weakly similar to BACN32G11.d [D.m	1.14	2.12
	115652	AA405098	Hs.38178	ESTs	0.82	4.67
	115875	AA433943	Hs.43946	ESTs; Weakly similar to Weak similarity	1.2	1.98
	116004	AA449122	Hs.76086	ESTs; Highly similar to small zinc finge	0.96	1.31
	116121	AA459254	Hs.48855	ESTs	0.97	1.55
75	116129	AA459956	Hs.49163	ESTs; Highly similar to putative ribonuc	1.08	2.73
	116190	AA464963	Hs.67776	ESTs	0.8	1.57
	116312	AA490494	Hs.65403	ESTs	1.37	2.65
	116732	F13779	Hs.165909	ESTs	0.92	1.8
	117602	N35020	Hs.44685	ESTs; Weakly similar to GOLATH PROTEIN	1.15	1.84
80	117950	N51394	Hs.75478	KIAA0956 protein	1.04	2.36
	117992	N52000	Hs.172089	Homo sapiens mRNA; cDNA DKFZp586B0222 (f	0.62	1.29
	118785	N75386	Hs.111867	GLI-Kruppel family member GLI2	1	1
	119717	W69134	Hs.57987	ESTs	1	1.4
	119814	W74069	Hs.58350	ESTs	0.78	1.77
85	120128	Z38499	Hs.91448	MKP-1 like protein tyrosine phosphatase	0.86	1.46
	120242	Z98443	Hs.86366	ESTs	0.83	2.01

	120483	AA252994	Hs.1578	apoptosis inhibitor 4 (survivin)	0.74	1.64
	121054	AA398504	Hs.97387	ESTs	1.05	1.93
	121326	AA404246	Hs.97031	ESTs; Weakly similar to Similar to phyto	0.98	1.3
5	121376	AA405699	Hs.166232	ESTs; Moderately similar to SODIUM- AND	0.91	1.83
	121457	AA411448	Hs.208985	ESTs	0.91	1.59
	121780	AA422086	Hs.124660	ESTs	0.46	0.55
	121781	AA422150	Hs.98370	cytochrome P540 family member predicted	1.07	1.54
	121844	AA425732	Hs.98485	gap junction protein; beta 2; 26kD (conn	0.94	1.4
	122059	AA431737	Hs.98749	EST	1.93	2.33
10	122338	AA443311	Hs.98998	ESTs	1	1
	122354	AA443772	Hs.186692	ESTs	0.88	1.39
	122591	AA453265	Hs.99311	ESTs; Weakly similar to MRJ [Hsapiens]	2.28	2.93
	122790	AA460156	Hs.99556	ESTs	0.88	1.3
15	123398	AA521265	Hs.105514	ESTs	1	1.93
	123518	AA608531	Hs.170313	ESTs	1	1
	123673	AA609471	Hs.112712	ESTs	1	1.15
	124000	D57317	Hs.74861	activated RNA polymerase II transcriptio	0.74	1.12
	124367	N24006	Hs.99348	distal-less homeo box 5	0.67	1.1
20	124447	N48000	Hs.140945	Homo sapiens mRNA; cDNA DKFZp586L141 (fr	1.19	1.7
	125756	W25498	Hs.81634	ATP synthase; H+ transporting; mitochond	0.93	1.59
	125769	AA382972	Hs.82128	5T4 oncofetal trophoblast glycoprotein	1.65	6.76
	125852	H09290	Hs.76550	Homo sapiens mRNA; cDNA DKFZp564B1264 (f	0.72	2.26
	125924	AA526849	Hs.82109	syndecan 1	1.22	2.25
25	126037	M85772	Hs.6066	KIAA1112 protein	1.36	1.63
	126214	N29455	Hs.74316	desmoplakin (DPI; DPL)	1.93	3.55
	126414	N78770	Hs.223439	ESTs	1.21	1.66
	126737	AA488132	Hs.62741	ESTs	1	1
	126743	AA179253	Hs.172182	pcl(A)-binding protein; cytoplasmic 1	1.3	2.16
30	126926	AA179546	Hs.832	ESTs; Highly similar to INTEGRIN BETA-8	2.53	2.8
	127432	AA501734	Hs.170311	heterogeneous nuclear ribonucleoprotein	1.57	2.12
	128218	H02682	Hs.99189	ESTs; Moderately similar to recombinatio	1.24	2.09
	128527	M31523	Hs.101047	transcription factor 3 (E2A immunoglobul	1.08	1.78
	128568	X60573	Hs.247568	adenylate kinase 3	1.23	3.48
35	128584	M11433	Hs.101850	retinol-binding protein 1; cellular	0.87	2.42
	128628	C14037	Hs.251978	EST	1.22	1.9
	128691	W27939	Hs.103834	ESTs	1.1	1.73
	128714	V00599	Hs.179661	Homo sapiens clone 24703 beta-tubulin mR	0.92	1.17
40	128733	AA328993	Hs.104558	ESTs	1.34	1.94
	128781	X85372	Hs.105465	small nuclear ribonucleoprotein polypept	0.9	1.34
	129052	AA496297	Hs.182740	ribosomal protein S11	2.59	3.19
	129095	L12350	Hs.108623	thrombospondin 2	1.04	3.2
	129241	AA435665	Hs.109706	ESTs; Moderately similar to HN1 [M.muscu	0.95	1.61
45	129665	M88458	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic relic	1.28	2.63
	129703	AA401348	Hs.179999	ESTs	0.97	1.63
	129720	AA476582	Hs.12152	ESTs; Moderately similar to SIGNAL RECOG	1.09	1.79
	129850	N20593	Hs.56845	GDP dissociation inhibitor 2	0.74	1.68
	129896	AA043021	Hs.13225	UDP-Gal:beta-GlcNAc beta 1,4- galactosyl	1.43	4.19
50	130069	AA055896	Hs.146428	collagen; type V; alpha 1	1.17	1.98
	130405	H88359	Hs.155396	nuclear factor (erythroid-derived 2)-lik	1.26	1.79
	130541	X05608	Hs.211584	neurofilament; light polypeptide (68kD)	1	1
	130599	M91670	Hs.174070	ubiquitin carrier protein	1.07	1.66
	130867	J04093	Hs.2056	UDP glycosyltransferase 1	1	4.8
55	131009	AA063596	Hs.22142	ESTs; Weakly similar to NADH-CYTOCHROME	0.93	1.05
	131028	U02040	Hs.2227	CCAAT/enhancer binding protein (C/EBP);	1	1.23
	131083	U66661	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.1	1.8
	131091	T35341	Hs.22880	ESTs; Highly similar to dipeptidyl pepi	1.28	1.98
	131144	C14412	Hs.23528	ESTs; Highly similar to HSPC038 protein	1.43	2.06
60	131148	C00038	Hs.23579	ESTs	0.88	3.38
	131164	Y00503	Hs.182265	keratin 19	1.19	2.77
	131185	M25753	Hs.23960	cyclin B1	0.86	3.84
	131219	C00476	Hs.24395	small inducible cytokine subfamily B (Cy	0.66	2.96
	131454	AA455896	Hs.2699	glypican 1	0.99	1.54
	131687	L11066	Hs.3069	heat shock 70kD protein 9B (mortalin-2)	1	1.18
65	131689	AA599653	Hs.30696	transcription factor-like 5 (basic helix	1	1.95
	131692	D50914	Hs.30736	KIAA0124 protein	1.55	2.39
	131786	AA135554	Hs.32125	ESTs	1	1.33
	131843	AA195893	Hs.184062	ESTs; Moderately similar to putative Rab	0.83	1.63
	131860	U02082	Hs.334	Oncogene TIM	1.08	2.2
70	131884	H90124	Hs.3463	ribosomal protein S23	1.23	1.24
	131903	AA481723	Hs.3436	deleted in oral cancer (mouse; homolog)	0.91	1.18
	131945	M87339	Hs.35120	replication factor C (activator 1) 4 (37	1	2.8
	131958	AA093998	Hs.3566	ESTs; Highly similar to phosphorylation	0.87	1.36
	131964	W42508	Hs.3593	ESTs	1	1.25
75	132001	J00277	Hs.37003	v-Ha-ras Harvey rat sarcoma viral oncoge	1.12	1.43
	132040	AA146843	Hs.172894	BH3 interacting domain death agonist	1	1.55
	132065	D82226	Hs.211594	proteasome (prosome; macropain) 26S subu	0.89	1.27
	132109	AA599801	Hs.40098	ESTs	1	1.05
	132112	AA150661	Hs.40154	jumonji (mouse) homolog	0.99	1.44
80	132123	AA447123	Hs.250705	ESTs	1.06	2.46
	132162	H89551	Hs.41241	ESTs	1.08	2.46
	132180	AA405569	Hs.418	fibroblast activation protein; alpha; se	1.02	4.56
	132309	AA460917	Hs.2780	jun D proto-oncogene	1.16	1.8
	132371	AA235448	Hs.46677	ESTs	0.8	1.26
85	132618	AA253330	Hs.5344	adaptor-related protein complex 1; gamma	0.5	1.49
	132736	U68019	Hs.211578	MAD (mothers against decapentaplegic; Dr	1.21	1.81

5	132771	AA488432	Hs.56407	phosphoserine phosphatase	1	1.3
	132833	U78525	Hs.57783	eukaryotic translation initiation factor	0.91	1.43
	132922	T23641	Hs.60668	KIAA1112 protein	1.16	1.53
	132959	AA028103	Hs.61472	ESTs; Weakly similar to unknown [S.cerev	1.02	1.88
	132994	AA505133	Hs.7594	solute carrier family 2 (facilitated glu	0.72	2.97
	133005	C21400	Hs.103329	KIAA0970 protein	0.88	1.34
	133065	X62535	Hs.172690	diacylglycerol kinase; alpha (80kD)	0.93	1.23
	133083	N70633	Hs.6458	chaperonin containing TCP1; subunit 2 (b	1.14	1.76
10	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso	0.97	1.43
	133134	T89703	Hs.65648	RNA binding motif protein 8	1.1	1.8
	133195	AA350744	Hs.181409	KIAA1007 protein	2.29	2.69
	133313	AA249427	Hs.70704	ESTs	1.07	1.68
	133331	T62039	Hs.158675	ribosomal protein L14	0.85	1.18
15	133438	D13370	Hs.73722	APEX nuclease (multifunctional DNA repa	0.91	1.45
	133445	T99303	Hs.73797	guanine nucleotide binding protein (G pr	0.94	1.68
	133483	X52426	Hs.74070	keratin 13	0.85	1.14
	133492	L40397	Hs.74137	transmembrane trafficking protein	1.1	1.69
	133504	W95070	Hs.74316	desmoplakin (DPI; DPII)	0.7	6.21
20	133517	X52947	Hs.74471	gap junction protein; alpha 1; 43kD (con	0.95	1.3
	133540	D78151	Hs.74619	proteasome (prosome; macropain) 26S subu	0.91	1.25
	133594	L07758	Hs.172589	nuclear phosphoprotein similar to S. cer	0.84	1.29
	133627	U09587	Hs.75280	glycyl-tRNA synthetase	1.09	1.99
	133671	T25747	Hs.75471	zinc finger protein 146	1.02	1.5
25	133859	U86782	Hs.178761	26S proteasome-associated pad1 homolog	1.11	3.33
	133865	F09315	Hs.170290	discs; large (Drosophila) homolog 5	1.84	6.7
	133913	W84712	Hs.7753	calumenin	1.15	1.86
	133963	L34587	Hs.184693	transcription elongation factor B (SIII)	1.3	1.91
	133982	U47621	Hs.207251	nucleolar autoantigen (55kD) similar to	1.3	1.99
30	134100	L07540	Hs.171075	replication factor C (activator 1) 5 (36	0.72	1.65
	134110	U41060	Hs.79136	LIV-1 protein; estrogen regulated	1.04	1.62
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	1	1.55
	134161	U97188	Hs.79440	IGF-II mRNA-binding protein 3	0.82	1.95
	134193	F09570	Hs.7980	ESTs	0.98	1.48
35	134367	X54199	Hs.82285	phosphoribosylglycinamide formyltransfer	1	2.8
	134402	U25165	Hs.82712	fragile X mental retardation; autosomal	1.26	2
	134457	D85963	Hs.174044	dishevelled 3 (homologous to Drosophila	1	1.47
	134469	X17567	Hs.83753	small nuclear ribonucleoprotein polypept	0.94	1.57
	134498	M53180	Hs.84131	threonyl-tRNA synthetase	1.2	2.64
40	134501	W84870	Hs.211568	eukaryotic translation initiation factor	0.84	1.36
	134507	M53488	Hs.84318	replication protein A1 (70kD)	1.7	2.93
	134548	U41515	Hs.85215	Deleted in split-hand/split-foot 1 regio	1.46	2.73
	134599	X99226	Hs.86297	Fanconi anemia; complementation group A	1.36	2.22
	134692	R73567	Hs.8850	a disintegrin and metalloproteinase doma	0.77	1.64
45	134693	N70361	Hs.8854	ESTs	1.09	1.82
	134806	Z49099	Hs.89718	spermine synthase	0.98	1.35
	134821	Z34974	Hs.198382	plakophilin 1 (ectodermal dysplasia/skin	0.99	1.4
	134864	Y08999	Hs.90370	actin related protein 2/3 complex; subun	0.95	1.42
	134914	U29615	Hs.91093	chitinase 1 (chitotriosidase)	1.18	1.29
50	134953	L10678	Hs.91747	profilin 2	0.95	1.76
	134993	AA282343	Hs.9242	purine-rich element binding protein B	0.98	1.73
	135051	C15324	Hs.93668	ESTs	1.35	2.11
	135158	U51711		Human desmocollin-2 mRNA; 3' UTR	0.86	1.16
55	Table 1B shows the accession numbers for those pkeys in Table 1A lacking unigenal D's. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the Accession column.					
60	Pkey:	Unique Eos probeset identifier number				
	CAT number:	Gene cluster number				
	Accession:	Genbank accession numbers				
	Pkey	CAT	Accessions			
65	100661	23182_1	BE623001 L05096 AA383604 AW966416 N53295 AA460213 AW571519 AA603655			
	100667	26401_3	L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700			
			AW605203 BE069721 AW382138 AW803776 BE463954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 A1632506 T29066			
			A1783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153			
70			BE546944 T69231 AW377441 AA907406 H50799 AW051416 AJ420712 BE620922 AJ279161 AA992549 W47198 BE005241 AJ342696 H50700			
			A1969974 AJ863855 AA374490 AW130675 AJ950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 AJ673068 AJ887890 AW804171			
			AJ675961 AW804172 AA778841 AL048050 AJ127757 AJ095568 AW204965 AW468978 W31898 AJ052595 AJ278771 BE464018 AJ081503 AJ824196			
			AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 AJ280942 T27619 BE621435 N66010 AW589527 AJ160414			
			AA283090 AA962536 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE565671 C00444			
			AA054555			
75	100668	26401_3	L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700			
			AW605203 BE069721 AW382138 AW803776 BE463954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 A1632506 T29066			
			A1783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153			
			BE546944 T69231 AW377441 AA907406 H50799 AW051416 AJ420712 BE620922 AJ279161 AA992549 W47198 BE005241 AJ342696 H50700			
80			A1969974 AJ863855 AA374490 AW130675 AJ950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 AJ673068 AJ887890 AW804171			
			AJ675961 AW804172 AA778841 AL048050 AJ127757 AJ095568 AW204965 AW468978 W31898 AJ052595 AJ278771 BE464018 AJ081503 AJ824196			
			AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 AJ280942 T27619 BE621435 N66010 AW589527 AJ160414			
			AA283090 AA962536 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE565671 C00444			
			AA054555			
85	101332	25130_1	J04088 NM_001067 AF071747 AJ011741 N85424 AL042407 AA218572 BE296748 BE083981 AL040877 AW499918 AW675045 H17813			
			BE081283 AA670403 AW504327 BE094229 AA104024 AJ471482 AJ970337 AA737616 A1827444 AW003286 AJ742333 AJ344044 AJ765634			

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

100780 458_127

100830 4002_1

100906 4312_1

100930 16865_1

102221 3851_1

101809 32963_1

102590 15932_1

101977 29073_1

102781 20812_1

A1948838 AW255336 AW172827 AA095289 BE046383 A1734240 W16699 A1660329 A1289433 AA933778 AW469242 AA468838 AA806983
 AA625873 W78031 BE206307 AA550803 A1743147 A1990075 AA948274 AA129533 A1635399 AA605313 A1624669 AW594319 A1624669 A1337434
 AA307705 BE550282 A1760467 A1630636 A121521 AW674314 AW076889 A1933732 A1686969 A1186928 AW074595 A127486 A1076644
 A1910815 H17814 AA310903 AW137854 T19279 AA026682 AA306035 AW383390 AW383389 AW383422 AW383427 AW383393 H08977
 AA306247 AA352501 AW403639 F05421 AA224473 AA305321 H93904 AA089612 AW391543 AW402915 AW173382 AW402701 AW403113
 R94438 N73126 H93466 AA090928 AA095051 T29025 AW951071 L47277 L47276 A1375913 BE384156 W24652 AA746288 AA568223 BE090591
 H93033 N57027 AA504348 AA327653 AW959913 N53767 AA843715 A1453437 AW263710 A1076594 AA583483 AW873194 AW575166 A1128799
 A1803319 AL042776 AW074313 A1887722 A1032284 AA447521 A1123885 N29334 A1354911 AW090687 AA236763 AA435535 AA236910
 AA047124 AA236734 AW514610 H93467 AA962007 A1445783 AA127259 A1613495 A1686720 A1587374 AA935731 AA702453 A1859757
 AA216786 A1251819 A1469227 AA805022 A1092324 N71868 AA958782 AA236919 AA809450 AA227220 AA765284 A1192007 AA768810
 AA805794 AA729280 AA806238 AW768817 N71879 A1050686 AA505822 AA668974 A1688160 BE045915 AW466315 AA731314 AA649568
 AA834316 AW591901 AW063876 AW294770 A1300266 A1336094 A1560380 AA721755 H09978 D20305 D29155 AW821790 A1150864 F01675
 A1457474 AW466316 AA550969 AA630788
 BE561958 BE561728 BE397612 BE514391 BE269037 BE514207 BE562381 BE514256 BE514403 BE514250 BE397832 BE269598 BE559865
 BE396081 BE560031 BE514199 BE560037 BE5600454
 AC004770 W05005 AA356068 AA094281 H29358 T56781 AW875313 L37374 BE312466 BE311755 BE207106 BE293320 BE018115 AW239090
 BE548830 AW247547 AA770682 BE397382 AA486713 T10111 T09340 AW498981 BE547280 AA356003 AW581520 AW875331 AA580720
 AW875336 BE276873 BE408229 AW188148 BE255166 BE253761 AW793727 AW373141 AW581548 AA471223 AA305950 BE263976 AA626820
 BE257408 AW360962 AA090655 C00312 BE312741 BE407213 AA203532 AW298199 AW248553 AW297794 AW731722 BE300586 AW731972
 AW615446 BE301599 AW615520 AA486714 AW440257 AA195516 AA564630 AA618079 AW192592 AW474955 AA604580 A1627461 AW765440
 A1680334 AL135548 A1683224 A1581126 AW245096 AW194154 H29274 N70363 AA629758 AA580602 AA862006 A183841 A1097667 A1928583
 A1358774 BE243487 AA620553 AA653297 AA292690 T10110 T238906 AA908544 AA340930 A1185438 T03328 T28844 A1687010 A1864965
 A1872575 BE388740 T56780 AW373138 BE258717 AA699671
 A1076916 BE298110 AW293955 AW672700 NM_003875 U10860 AW651755 BE297958 C03806 A1795876 AA644165 T36030 AW392852
 AA446421 AW881866 A1469428 BE548103 T96204 R94457 N78225 A1564549 AW004984 AW780423 AW675448 AW871454 AA305698
 AA879433 AA535069 A1394371 AA928053 A1378367 N59764 A1364000 A431285 T81090 AW674657 AW674987 AA897396 AW673412 BE063175
 AW674408 A1202011 R00723 A1753769 A1460161 AW079585 AW257544 A1873729 D25791 BE537646 T81139 R00722
 J04129 NM_002571 AA293088 AA477016 AA404631 T28299 AA476904 AA433955 AA430486 AA495907 A1151391 AA291495 AA402723 W25651
 A1076816 A1826712 AW296294 AA239479 A1276581 AW044154 A1080180 A1417985 A1274168 A1474212 AA495908 AA635664 A1092114
 A1804952 AA479874 A1597661 A1420511 AA479738 AA421417 AA421247 AA436220 AL047797 M34046 N42277 AA228076 W02698 AA20297
 AA434011 A1369971 AA479731 A1865541 A1418020 AA421246 AA452764 AL048051
 NM_006769 U24576 AW161951 AW160473 AW160465 AW160472 AW161059 A1824831 AW162635 A1990366 AW162477 AW162571 A1520836
 AW162352 AW162351 AW162752 A1962216 A1537346 AA853902 H17667 BE045346 BE559802 BE255391 AA985217 AA236501 A1297057
 AW366451 T34489 D56106 D56351 A1936579 AW023219 AW889335 AW889120 AW889232 AW889175 BE093702 AW889349 A1417546
 A195298 AA912579 A143356 AW902211 R64717 AW157236 A1815242 D45274 AW263991 AA442920 AA129965 AL035713 A1923255 A1949082
 A142826 A1684160 A1701987 A1678954 A1827349 BE463635 AW628092 AW302281 AA493203 BE348856 BE356419 AW193969 AW673561
 AW592609 A1224044 H43943 AA091912 L49632 R48353 A1568409 R48256 A1199046 H27986 H43899 A1678759 A1680310 A1624220 H17052
 AA156410 N56062 A1699430 AA664529 T09406 T10459 AA627506 A1379584 N63831 N88633 AW022651 AA971281 AA248035 A1039197
 A1914689 AA973825 AL047305 AA129966 A1798369 AW264348 A1445879 A1658759 N67924 A1933507 A1216121 A1333174 T10972 A1375028
 A1186756 A1273778 AA610487 A1797946 AA853903 AA903939 A1338587 A1278494 AW627595 AA904019
 M58849 AA315280 NM_004004 AA315269 BE142653 AA641400 AW802042 BE152893 AW383155 AA490688 AW117930 AW384563 AW384544
 AW384566 AW378307 AW378323 AW839085 AA257102 AW378317 AW276060 AW271245 AW378298 AW384497 A1598114 AW254544 A1018136
 AW021810 AA961504 AW086214 AW771489 AW192483 A1290266 AW192488 AW384490 AW007451 AW890895 AA544660 AA613715
 AW020066 A1783695 A1589498 A1917637 AW254471 AW384491 A1816732 AW368530 AW368521 AW358463 AA461087 A1341438 A1970613
 A1040737 A1418400 AA947181 AA962716 A1280695 AW769275 AW023591 A1160977 AA055400 N71882 AA490466 AW243772 AW316636
 A1076854 AW511702 N69323 H88912 AA257017 A1952506 H88913 A1912481 AA600714 BE465701 N64149 C00523 N64240 AA677120
 R61573 BE050529 X98091 AA297307 BE537267 BE566138 BE566139 F11561 BE564795 BE568776 AW064005 BE566479 BE380035 BE567012
 BE568634 BE566568 AA298050 BE566043 BE568813 BE568618 AA283070 BE565414 BE566738 BE568585 BE566567 BE566116 BE566433
 U82136 AF049140 BE567057 BE567742 BE567403 BE564316 BE567400 BE568854 BE566588 AA448772 AA071363 AW732642 BE564996
 AA297763 AA278550 AA421083 AA298184 AA091007 AA984577 AA205916 N28759 AL031291 C15757 C15761 H02778 BE566410 A1129335
 AA419499 N87741 BE379689 BE004824 BE379611 D25874 AA148454 AA323654 AW950311 AA448795 AW749423 AA773386 AA773843
 AW020327 BE348580 BE504258 BE549990 BE220200 A1673334 A1020259 AA975516 D61421 A1168688 AA102843 AW246621 A1276203
 A1074054 A1633824 A1962927 A148928 N50969 A1308911 AA410994 AW373025 AA148455 H02620 AA688293 A1246318 N22220 A1917777
 A1050943 A1097286 AA663794 AW368662 AW627826 AW087874 A1253060 AA749154 AA832236 A1192358 AW204676 AA448676 AA764891
 BE439467 AA661534 AA258061 A1090546 AA995157 A1051011 AA584421 A1026032 AW591338 AW589563 AA776914 AW024684 AA421002
 F09219 BE464500 A1383595 AA954244 AA601583 AA737304 AA195549 AA805778 A1055876 AA164942 AW013861 A1672608 AW517211 D59441
 AW582574 AA160935 BE566501 BE564612 BE565353 BE568195 BE565447 BE568302 BE566097 BE565470 BE564249 AL036217 AW749424
 BE567424 AA102842 AA314761 AV661237 C14211 AA651866 AW798997 AA470805
 AF11219 AL050318 T24804 AW248136 BE386341 BE263177 W16677 BE250224 BE563669 BE267405 BE546577 AV651354 AV651292
 A1346903 A1539128 A1189171 S83364 AW073849 A1816760 AW073309 A1422590 AA296692 A1860301 A1805446 N77735 A1340328 BE092630
 AW028742 BE088442 AA657742 AA742438 AW170086 A1038920 A1432379 N36073 A1936194 AA868655 AA836112 A1077505 BE080433
 A1375014 A1126547 A1348244 A1346077 A1748952 N26915 A1753574 A1093341 A1278762 BE092517 N74204 H08158 T58149 A1129303 N58366
 AA524456 BE122661 AA542925 A1246120 A1735203 AA706829 AA877544 A1082289 AA926687 N92840 AW249798 AA934763 AW998363
 A1128532 N25202 A1240209 AW118892 N80744 R35655 A1342321 A1340141 AW878792 A1857321 H09610 W04601 AW006650 AA126006
 AA553675 A1052791 AW059835 A1041906 AA814658 AW002059 AA729483 A1609301 AA994633 AA903651 A1459183 T95072 AW088630
 AA126112 A1800091 A1561215 H17502 AW475072 A1819003 A1683272 A1262701 AW793140 T81787 R95586 A1275160 A1310420 A1698928
 AA159174 A1827968 F30305 F30309 AA806662 A1091923 AW878722 AA583430 AW571913 A1674584 AA292533 A1079471 AA642325 AA719050
 AW793172 AA305476 AW103745 T23459 N79525 A1784438 AA534551 AW193751 A1074380 BE281214 T32229 W25066 W01205 T63086
 AW795348 A1361287 AW795353 AW795349 AA594759 A1400295 D11489 A1370689 AA482356 AA485295 W40151 AA564661 AW300745
 A1346938 A1374975 A1423782 AW193899 AA612604 A1183409 AA996156 AW366963 AW366977 A1284860 AA846503 A1985064 AA844576
 AA737921 AA873274 BE241546 BE241540 AA484058 AW468970 AA127876 AA159120 AW001568 AW795213 AW795258 AW795330 BE250589
 BE387572 AA910895 AA161217 BE250360 W31500 T95167 A1719306 A1359224
 BE258778 BE281230 BE410044 T37323 AW672694 AW410439 NM_006429 AF026292 T35505 BE542333 T08940 AU076737 AW247471
 BE393215 AW328640 BE542408 T32170 BE302544 T31955 BE206898 BE275738 T32570 BE386426 BE298746 BE389937 BE293991 BE315289
 BE389578 R34739 R15312 BE279365 BE277756 AL036019 T37325 BE277779 BE302962 AL047294 BE276505 T09070 T33673 BE312580
 AW387774 BE257175 AW674367 BE253331 BE270344 BE299831 BE273576 T32062 A1751831 BE618381 AA304895 BE252268 U45364
 BE256790 BE207199 BE256209 BE251941 BE250791 BE313955 BE269806 BE543623 BE279212 BE252289 T31699 BE262220 T31669
 AA315781 AA192212 N84547 BE292737 BE259631 AA232179 A1133144 T31292 AA315945 BE407301 BE251184 BE409006 A1880158 A1904003
 A1904114 AW651788 AW651763 R58247 BE271897 U83843 C05298 BE261609 BE255973 AA351650 N84631 BE263637 AW452919 AA328465
 AA324549 AW578525 BE252296 BE257551 AL048332 BE208630 AA359336 AW327897 AA151742 AA305816 BE076862 BE076796 BE263161
 AA323785 AA676588 AA626555 AA078917 W87657 R09002 R94021 AA312032 BE276665 AA295608 AW407162 AA329374 AW877912 N27885
 AA369255 AA360968 BE250476 N85427 BE265569 A1278639 A1816576 A1691037 AW328583 A1567949 A1983455 A1927732 A1811297 A1571508
 AW073674 BE296039 BE467326 A1828795 A1816578 AW511604 A1921213 AW152427 A1795787 A1801618 AW168866 A1628144 A1890339
 AW173690 AW511540 BE535620 AA383014 BE301164 A1866596 AW514909 AA658050 AW575243 AA074631 A1093488 AW575408 AW675443
 AW615636 AW732207 AW377638 AA321784 AA641629 AA633105 AA527640 AW129146 AW615672 BE394607 AA483902 AA475032 BE378532

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

AA872808 AI469388 AW105268 BE047301 AW591843 AW410066 AW517153 AI950495 AA746641 AI914878 AA873185 AI696911 AA548625
AA911505 AA148762 AW674535 AI587329 BE328328 AW270348 AA158225 AW117705 AW474997 AW519193 AA614757 AW664383 AI082647
AW590973 AI476711 AA192213 N88741 BE464552 AW072679 AI453708 AA152166 AA805924 AI581078 AI125768 AW173484 AI961980
BE300766 AI199598 AI636792 AW247333 AW272861 AA078818 AA150012 AA551232 AA678821 AW873869 AW768266 AI660315 AA319210
AA814551 AA157994 AA318886 AI582962 AW089224 AI356098 AI343694 AW072598 N12054 AI301249 AA742924 H17917 AW328584
AW248898 AI751830 AA907816 R08898 AW087989 AI828300 AA148596 AI269577 T33426 AA213571 AI973201 AA666279 R49612 AI573183
AW797962 AW410068 AW769666 AI962097 AI475204 D57490 AW517531 BE245270 AW470008 T33427 AW005731 AI795795 AW272981
T15747 AA552875 T23644 AW361289 AI758558 BE207435 AA876958 T03361 AA883569 F37533 AA582321 AW082524 R42212 AA973847
T18900 AA085202 AI558967 AI302418 AA948667 AA745670 T08939 T33724 T33722 BE621568 D57489 D25906 BE621151 F16510 C05666
T35127 AA630427 AI933481 AA309426 AI918440 BE561854 BE618866 BE394675 BE296173 AW951687 BE383739 BE616141 BE312730
BE535351 AW080575 BE313330 BE616664 AI354390 AA847315 BE544509 BE515212 BE297833 BE278808 BE544844 AW090178 AI890664
BE546708 AW189943 BE274412 BE382399 BE266392 BE254949 BE280696 BE383237 BE261756 BE257721 BE312683 BE275476 BE514880
BE545314 BE313587 BE384537 BE386691 BE264813 AW592575 AI336332 AI278541 AI795791 BE222662 AW249316 AA314361 AI036012
AW402923 BE266845 AA075945 AA314436 BE384640 AW731769 AW957077 AA552234 AA573560 AW367038 AA313399 AI983873 BE410159
BE263803 BE514339 BE409073 BE281296 BE543396 BE395387 BE088360 BE546946 BE546570 BE390626 AA074638 AA301821 AA4854230
AW582379 AI949222 AW029572 AA515843 AW272394 BE250234
C14322 W74050 AI074232 AA595624 BE048956 AI148417 AI583145 AI473460 AI801688 AW573593 AI950741 AI628140 AW467921 R38105
AI149258 AI247584 AI078378 AI139850 AA489411 W24744 R98104 AI033826 AA699589 AI033120 N55544 W89894 AW970771 AA703362
AA099138 AA706792 AA046150 H98981 AI915674 AA953018 AI972749 AI921343 AA909044 AA094751 AI203124 AA582143 AI466654
AW235415 R70377 AA099236 F20703 AA524436 R69484
H04043 D60988 D60337
AI204246 AI204250 AI194050
H83265 T63524 AA304359 AW980551 AI672874 AI749427 AA227777 AW027055 AA971834 T49644 T54122 AI983239 AI808233 T91264 T96544
AI350945 AI079114 R72382 T48788 R48726 W385418 AI095484 T49645 AA928653 AA570082 AW007545 T57178 AA516413 AA913118 T57112
AA564433 AA774503 AA367671 T59757
W78816 AI720806 AI633854 AI632086 AI666663 N70894 AW571809 AI383592 AI201348 W80715 N91880 AW963101 AA339011
AB033023 BE391906 BE275965 BE277872 BE003882 AA313774 BE019159 BE298024 BE299727 BE300011 BE380277 BE394764 N87550
BE409419 BE408652 BE408197 AI119932 AA822427 AI816265 AA610118 T07318 AA019839 AA634430 BE205794 BE049461 AA042322
AI652711 AI917645 AA630045 AW191969 AI817882 T17271 AI803663 AI095533 H46019 AW592438 AI624835 AI675552 D51149 AW132058
AA639614 AI925762 AW088153 T17455 AA018640 AW751475 BE300241 AI816255 BE391981 AW408671 AA353910 AW875446 AW5467503
AW875926 AW875645 AW875647 AW938037 AL138042 AW892619 BE243018 AW995454 BE246381 BE009082 BE278921 AW957842 AA262454
H30121
W72062 AF088057 W76255 AI827219 AI631461 AW449295 AI354957 AI913803 T62772 AI222040 T62921 T63781
AI678765 H12175 R14664 AI914049 AA995383 H08009 H19418 AW953728 AI358021 AA587361 AI269377 AA369905 AW957113 H27693
AI300474 H73776 W74397 AA579604 AI131018 W72331 AI719085 AA568348 AI859045 AI814819 AI888714 BE467470 AW131268 H19419
H27694 AI342165 AI914155 AA534872 BE018176 R60206 H11647 R45641 AI860466 BE301656 AI125453 AI498120 AA593735 AA879110
AI016404 T35018 AA588397 AW449767 AA470365 BE501139 AA588354 AI337500 AW078532 Z41279 AI125449 AA935725 AA040338
L42583 NM_005554 L42601 BE183076 AI541221 BE140567 L42610 V01516 J00269 AW275792 AW383052 AW380143 AI541102 BE612846
AI541344 AW238368 BE613405 BE615705 BE615530 BE615301 AW379823 AW794706 AA194806 AA194992 AW384024 AW384000 AA541239
AI246504 AI540333 AW238681 AA640939 AI540863 AI608860 AW662564 AW366725 AW368983 AW366870 AA596020 AW794721 AW794511
AI591181 BE182523 AW794644 AW794620 AI935234 AI608903 AI608623 AW797060 AW084935 BE182517 BE182319 AI890082 AW238346
AW797012 BE182522 AW794838 AI608794 AW304289 AA147193 AA595995 AW381128 AW366720 AA583718 AI828416 BE122864 AW368343
AA431080 AW082038 AW380976 AA587144 AA443636 AW872937 AW794448 AW378382 AW085761 AW794718 AW126395 AA583587
AA583991 AA583994 AA586886 AA586880 AW368365 AI814460 AA586991 AI282829 AW378406 AA586721 AI609242 AA431973 AA232959
AI831095 AW263854 AW378391 AW378415 AW378381 AA036990 AW238395 AI285446 BE208219 BE049526 AA583605 AA583918 AW366711
AI285580 AW082642 AI285712 AA582875 AW591216 AW368719 AW378408 BE122835 AA582976 BE350422 AA418328 AI541454 AI565930
AA583700 AA150575 AW238427 AI287474 AA912658 AA584223 AW238526 C17918 AW136169 AA159847 AI923797 AI609009 BE182479
AI915198 AW378114 AA147179 AA584239 AA150532 AW168862 AW085999 AW082480 AA659742 AW079703 AI872793 AA583981 AI824571
BE182316 BE182507 AA233331 AI824572 AI540586 D29492 BE182931 AA036948 BE551821 D29401 AW378365 C00141 D29181 D29567
AW103359 W95238 AI991663 AA587298 BE184608 AA099833 W95121 W95150 D29584 AI934111 D29456 D29533 AW265380 D29290
AW238463 AA121041 D29204 AA595925 D29441 AW081840 AA587018 D29323 AA582891 BE182433 BE182437 BE158295 BE182434
AW015534 AA314369 AA290715 BE568683 AW629494 D28364 AW995678
AI907114 AA580734 AL041945 AA101515 AA121344 D78130 NM_003129 AA341650 T84166 AF098865 AA130976 BE089553
T66122 AW175590 F05344 AI14790 R12900 AA194871 AA132258 D78129 AA132213 AW948930 AW948919 AA263053 AW948593 AW948940
AA278558 F058085 N26940 N40818 AW021255 AA054851 AA663379 AW948795 AW948893 AA400356 AW948911 N85024 W78844 AI341546
AI760182 AA268763 BE617763 BE617263 AW263690 BE049454 BE617928 AW515038 AW950584 AA601009 AI079194 AA147204 AW083163
AA130981 AI218369 AA604784 AI806257 AI595558 AA232318 AA258065 AI471982 AA687949 AI143944 N30172 AA400196 AI769049 AI084342
AI221380 AA948469 AI802469 H05720 AA113270 AA158138 AA076231 AI521024 AI810962 AI133616 AA805106 AA101516 R40052 R50778
R43280 T65036 AW131924 AA114261 AA152331 F09650 AA580614 AA558927 C75491 Z38352 AA954595 C75606 W80742
D56165 M36981 X58965 NM_002512 BE379177 AA314838 BE256445 BE252016 AW248343 AI720933 AW085701 BE386050 BE619742
BE277805 AA147951 AA603113 BE253293 AI246588 AI183405 AI954174 AI126891 AI829101 AI123832 AW129670 AA471268 AW170242 AW873079 AW148166 AI086820
AA482961 AI003658 H43261 AA657978 AI735072 R83138 AA722002 AA626271 AW273877 BE464626 AA071483 AA429973 AA494342
AA620436 AA775597 AA775601 AA826847 AI192585 AA826359 AA411159 AI193419 AI204013 AA705323 AA716255 AI784611 AI081144
AI128227 AA828464 AI148911 AI493446 AI626084 AI189180 AI721196 AI190618 AA284987 AI128543 AA632064 AI333073 AI278470 AA131688
AI491768 AA937581 AA630065 AA834257 AW249841 AA583742 AI309756 AA961676 AI760860 AA557818 AA954238 H43655 AI302564
AA127545 AI609219 H20426 AI042292 AI055466 AA581836 W47002 AA422057 AA937573 F29757 AA829208 AW327462 AA372098 W02144
AA036805 AA487365 AA961037 AI139946 AA487250 AA737118 AI952504 AI242293 AA605552 AI708401 AI633133 AA630848 AA654317 F24128
AI434165 W46252 AW043879 AI033763 F37228 AA687809 N49087 AA876981 AA506947 AI914572 AI833284 F22253 AA026222 R50166
AI219267 N27095 AA496512 AI784222 AI289904 AA513146 AA528547 AA418700 F36721 AI880700 AI601170 AI862851 AI708633 AA524499
AA642220 AA486628 AI718709 W80579 AI720547 F20718 AA649543 AA588229 N40603 H46029 BE262669 BE391069 BE537538 AI510761
AI906968 AI318611 H46099 AI472604 T60667 AA373087 W32479 AA514034 BE619183 AA134672 AA127544 H26942 BE536689 AW327461
AA422139 AW262357 AW327348 F33510 AI630382 AW827126 F27133 AI335189 AW517599 W80471 AA858514 N89681 BE393173 AA517760
AA584268 AA460537 AA446261 H20425 N64040 AW276801 AA316367 AA071232 BE545409 AA308292 BE274447 AA380861 AA340038
AA341806 AA865579 AI018634 AI766314 AI919302 AA872367 AA991404 AI906961 AA888375 BE621012 AA505388 AA935192 AA290828
R50220 H50814 H44721 AW951723 AA514796 AA418708 AW573377 AA379622 AA977995 AA708224 AA708216 AI318249 AI318233 AA411180
AA062221 AA316774 AA486808 AI500094 AA096362 AW583742 BE536422 BE618653 R70203 AA131732 AA345048 BE562720 T28342
NM_004415 AI031058 W77630 BE149760 AW752599 AW848723 AW376697 AW376817 AW376699 AW848371 AW376782 AW848789
AW361413 AW849074 AW997139 AW799304 AW799309 BE077020 BE077017 BE185187 AW997196 BE156621 BE179915 BE080661 BE143155
AW890985 BE002107 AW103521 AA857316 AW383133 BE011378 AW170253 BE185750 AW86475 BE160433 J05211 BE082576 BE082584
BE004047 AW607238 AW377700 AW377699 BE082526 BE082505 BE082507 BE082514 AW178000 AW177933 AI905935 AW747877 AW748114
BE148516 AW265328 AW847678 AW847688 AW365151 AW365148 AW365153 AW365156 AW365175 AW365157 AW365154 AW068840
BE005272 AW365145 BE001925 BE182166 BE144243 BE001923 AI951766 AI434518 BE184920 BE184933 AI284090 BE184941 AW804674
BE184924 C04715 W39488 AW995615 BE184948 BE158646 AW606653 AA099891 AA131128 AA337270 AA340777 AW384371 AA852212
R58704 AW366566 AW364859 AA025851 AA025852 AA455100 AA719958 AW352220 AW996245 BE165351 BE073467 AA377127 AW890264
AW609750 AW391912 AW849590 T87267 AW853812 AA852213 W71419 BE009090 AA056401 H91011 AW386529 AW390272 AW87

5

10

15

20

25

AW674920 N57176 AA026480 AW575767 H93284 AA026863 AW177787 AA026654 AW177786 BE092134 BE092136 AW177784
AJ022862 BE091653 AW376811 AW848592 AA040018 BE185331 BE182164 AA368554 AW951576 T29918 AA131077 W95048 W25458
AW205789 H90899 N29754 W32490 R20904 BE167181 BE167165 N84767 H27408 H30146 AJ190590 C03378 AJ554403 AJ205263 AA128470
AJ392926 AF139065 AW370813 AW370827 AW798417 AW798780 AW798883 AW798569 R33557 AA149190 C03029 AW177783 AA088866
AW370829 AA247685 BE002273 AJ760816 AJ39101 AW879451 AJ700963 AA451923 AJ340326 AJ590975 T48793 AJ568096 AJ142882 AA039975
AJ470146 AA946936 BE057737 BE067786 W19287 AA644381 AA702424 AJ417612 AJ306554 AJ686869 AJ568892 AW190555 AJ571075 AJ220573
AA056527 AJ471874 AJ304772 AW517828 AJ915596 AJ627383 AJ270345 AW021347 AW166807 AW105614 AJ346078 AA552300 W95070
AJ494069 AJ911702 AA149191 AA025864 AJ830049 AJ887258 AW780435 AJ910434 AJ819984 AJ858282 AJ078449 AJ025932 AJ860584 AJ635878
AA026047 AA703232 D12062 AW192085 AA658154 AW514597 AW591892 T87181 AA782066 AW243815 AW150038 AW268383 AW004633
AJ927207 AA782109 AW473233 AJ804485 AW169216 AJ572669 AA602182 AW015480 AW771865 AJ270027 AA961816 AA283207 AJ076962
AJ498487 AJ348053 AJ783914 H44405 AW799118 AA128330 AA515500 AA918281 W02156 AJ905927 AA022701 W38382 R20795 T77861
AW860878
100528 45979_1 BE386801 AU077299 AA143755 BE302747 AA853375 U30162 BE274163 BE277479 BE408180 BE274874 C15000 AA047476 T20799 AJ359165
AJ638794 AJ151283 AJ863925 AW444977 AJ207392 AA931263 AA443112 R40138 AW068538 AA351008 AA676972 R62503 AA916492 AW001865
H42334 H38280 AA121497 AA114137 AJ750938 M17783 AA383786 BE274462 AJ753182 C05975 AA347404 AW069298 AJ754351 AJ754044
AA188808 AA186879 AA565243 AL040655 AA456177 AJ750722 AA045756 AA213580 C16936 AW578747 AW753731 H41632 N44761 R58560
R61260 AA039902 N59721 AW992543 R68380 AA149586 T29017 H03739 BE383822 BE387105 BE408251 BE410425 H41560 AA247591
BE385677 AJ752233 AJ566195 AA868004 AJ424523 AW753720 AA852159 BE385803
100559 2260_1 NM_000094 L02870 D13694 S51236 M96984 AW946290 M65158 AJ285422 D29523 AL119886 AW630655 L06862 AJ884355 AW168737 T29085
AW797005 AW801340 AJ355504 AW079048 AW801337 AJ690455 AJ972063 AW268565 W68588 AA587326 AA883498 AJ033523 AW510356
AW591998 H98463 AL043852 AJ150055 AJ566239 AJ624803 AA844717 H40670 AA922334 AJ864424 AW615094 AW451233 AJ302203 F31221
AJ872170 W68589 AA904478 AJ917631 AW014208 AW450759 AA847625 AJ284033 AA848176 AA598507
100576 9986_1 X00356 NM_001741 M26095 X03662 M12657 X02330 X02330 AA716058 AW296074 X04861 AJ695720 AA719597
124357 genbank_N22401 N22401
101624 entrez_M55998 M55998
101625 entrez_M57293 M57293
135158 57953_1 AL037551 AJ804716 AW439811 AJ569470 AA075299 AJ738572 AJ270388 AJ816783 AW263026 AJ633951 AJ655285 AJ990572 AJ950425
AW241533 AA916883 AA576693 AA160156 AA613783 AW078884 AJ888282 AJ275241 AJ133467 AA164921

Tables 2A-8C were previously filed on November 9, 2001 in USSN 60/339,245 (18501-004100US)

Table 2A shows 504 genes down-regulated in lung tumors relative to normal lung and chronically diseased lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

	Pkey:	Unique Eos probeset identifier number										
	ExAccn:	Exemplar Accession number, Genbank accession number										
	UnigeneID:	Unigene number										
10	Unigene Title:	Unigene gene title										
	R1:	90th percentile of AI for normal lung samples divided by the 80th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.										
	R2:	median of AI for normal lung samples divided by 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.										
15	R3:	median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.										
	R4:	average of AI for normal lung samples divided by average AI for squamous cell carcinoma and adenocarcinoma lung tumors.										
	R5:	median of AI for normal lung samples divided by the 90th percentile of AI for adenocarcinomas.										
20	R6:	median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.										
	R7:	average of AI for normal lung samples divided by the 90th percentile of AI for squamous cell carcinomas.										
	R8:	median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for squamous cell carcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.										
25	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5	R6	R7	R8
	100095	Z97171	Hs.78454	myocilin; trabecular meshwork inducible	40.20							
	100115	NM_002084	Hs.336920	glutathione peroxidase 3 (plasma)								3.46
30	100138	U83508	Hs.2463	angiotensin 1			2.30					
	100299	D49493	Hs.2171	growth differentiation factor 10		11.00						
	100306	U86749	Hs.80598	transcription elongation factor A (SII);						3.06		
	100447	NM_014767	Hs.74583	KIAA0275 gene product								3.16
	100458	S74019	Hs.247979	Vpre-B	42.40							
	100862	AA005247	Hs.285754	Hepatocyte Growth Factor Receptor						4.13		
35	100959	AA359129	Hs.118127	actin; alpha; cardiac muscle				125.60				
	101032	BE205854	Hs.46039	phosphoglycerate mutase 2 (muscle)	36.40							
	101081	AF047347	Hs.4880	amyloid beta (A4) precursor protein-bind				34.60				
	101088	X70697	Hs.553	solute carrier family 6 (neurotransmitter				193.20				
	101125	AJ250562	Hs.82749	transmembrane 4 superfamily member 2						3.10		
40	101180	U11874	Hs.846	Interleukin 8 receptor; beta				54.86				
	101308	L41390		*Homo sapiens core 2 beta-1,6-N-acetylgl	33.20							
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do				36.40				
	101345	NM_005795	Hs.152175	Calcitonin receptor-like			2.29					
	101346	AJ738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				70.55				
45	101397	M26380	Hs.180878	lipoprotein lipase								3.54
	101414	NM_000066	Hs.38069	complement component 8; beta polypeptide							3.81	
	101435	NM_001100	Hs.1288	actin; alpha 1; skeletal muscle				34.60				
	101507	X16896	Hs.82112	interleukin 1 receptor; type 1				37.60				
50	101530	M29874	Hs.1360	cytochrome P450; subfamily IIB (phenobar			2.54					4.25
	101537	AJ469059	Hs.184915	zinc finger protein; Y-linked								
	101542	NM_000102	Hs.1363	cytochrome P450; subfamily XVII (steroid		5.50						
	101545	BE246154	Hs.154210	EDG1; endothelial differentiation, sphin	39.40							
	101554	BE207611	Hs.123078	thyroid stimulating hormone receptor		13.00						
55	101560	AW958272	Hs.83733	Interleukin 12 receptor; alpha 2, exon								3.38
	101574	M34182	Hs.158029	protein kinase; cAMP-dependent; catalyti						4.37		
	101605	M37984	Hs.118845	tropomyosin C; slow								3.80
	101621	BE391804	Hs.62661	guanylate binding protein 1; interferon-	30.20							
	101680	AA299330	Hs.1042	Sjogren syndrome antigen A1 (52kD; ribon						3.37	2.75	
60	101829	AW452398	Hs.129763	solute carrier family 8 (sodium/calcium								
	101842	M93221	Hs.75182	mannose receptor; C type 1				38.20				
	101961	AW004056	Hs.168357	*Hs-TBX2-T-box gene (T-box region) [huma			2.32					
	101994	T92248	Hs.2240	uteroglobin								6.85
	102020	AU077315	Hs.154970	transcription factor CP2			2.45					
65	102091	BE280901	Hs.83155	aldehyde dehydrogenase 7								6.75
	102112	AW025430	Hs.155591	forkhead box F1	54.60							
	102190	AA723157	Hs.73769	folate receptor 1 (adult)								3.98
	102202	NM_000507	Hs.574	fructose-bisphosphatase 1								3.62
	102241	NM_007351	Hs.268107	Multimerin			2.32					
	102310	U33839		Accession not listed in Genbank		7.00						
70	102397	U41898		*Human sodium cotransporter NKST1 mRNA,	29.40							
	102571	U60115	Hs.239069	*Homo sapiens skeletal muscle LIM-protein								3.75
	102620	AA976427	Hs.121513	*Human clone W2-6 mRNA from chromosome X						3.07		
	102636	U57092		*Human ataxia-telangiectasia locus prote			2.40					
75	102667	U70867	Hs.83974	solute carrier family 21 (prostaglandin			3.15					
	102675	U72512	Hs.7771	*Human B-cell receptor associated protel						3.56		
	102698	M18667	Hs.1867	progastrin (pepsinogen C)								4.51
	102727	U79251	Hs.99902	opioid-binding protein/cell adhesion mol					12.00			
	102852	V00571	Hs.75294	corticotropin releasing hormone	37.40							
80	103026	X54162	Hs.79386	thyroid and eye muscle autoantigen D1 (6					13.00			
	103028	X54380	Hs.74094	pregnancy-zone protein	28.80							
	103098	M86361		Human mRNA for T cell receptor; clone IG					10.00			
	103117	X63578	Hs.295449	parvalbumin		6.00						
	103241	X76223		Hsapiens MAL gene exon 4			2.47					
	103280	U84722	Hs.76206	Cadherin 5, VE-cadherin (vascular epithe			2.69					
85	103360	Y16791	Hs.73082	keratin; hair; acidic; 5							2.16	

PCT/US02/12476

96

WO 02/086443

PCT/US02/12476

	109001	AJ056548	Hs.72116	ESTs, Moderately similar to hedgehog-int	2.57				
	109003	AA147497	Hs.71825	ESTs				2.11	
	109004	AA156235	Hs.139077	EST	5.60				
5	109065	AA161125	Hs.252739	EST			10.00		
	109250	H83784	Hs.62113	ESTs; Weakly similar to PHOSPHATIDYLETHA				3.44	
	109490	AA233416	Hs.139202	ESTs				2.92	
	109510	AJ798863	Hs.87191	ESTs	2.40				
	109578	F02208	Hs.27214	ESTs	10.00				
10	109601	F02695	Hs.311662	EST			40.80		
	109613	H47315	Hs.27519	ESTs			54.40		
	109550	R31770	Hs.23540	ESTs	31.20				
	109682	H18017	Hs.22869	ESTs	8.40				
	109724	D58899	Hs.127842	ESTs			29.40		
15	109782	AB020644	Hs.14945	long fatty acyl-CoA synthetase 2 gene			8.00		
	109833	R79864	Hs.29889	ESTs	10.00				
	109837	H00656	Hs.29792	ESTs		6.49			
	109977	T64183	Hs.282982	ESTs				2.75	
	109984	AJ796320	Hs.10299	ESTs			107.00		
20	110146	H41324	Hs.31581	ESTs; Moderately similar to SYNTAXIN 1B				2.22	
	110271	H28995	Hs.31330	ESTs				3.48	
	110280	AW874263	Hs.32468	ESTs	44.20				
	110420	R93141	Hs.184261	ESTs			32.00		
	110578	T62507	Hs.11038	ESTs	28.40				
25	110634	R38905	Hs.35992	ESTs			20.00		
	110726	AW961818	Hs.24379	potassium voltage-gated channel; shaker-				4.15	
	110837	H03109	Hs.108920	ESTs; Weakly similar to semaphorin F [H.			56.80		
	110875	N35070	Hs.26401	tumor necrosis factor (ligand) superfamily		3.13			
	110894	R92356	Hs.66881	ESTs; Moderately similar to cytoplasmic	5.33				
30	110971	AJ760098	Hs.21411	ESTs			44.60		
	111023	AV655386	Hs.7645	ESTs	32.40				
	111057	T79639	Hs.14629	ESTs			17.14		
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f				4.58	
	111330	BE247767	Hs.18166	KIAA0870 protein				3.42	
35	111374	BE250726	Hs.283724	ESTs; Moderately similar to HYA22 [H.sap				3.91	
	111442	AW449573	Hs.181003	ESTs			33.20		
	111737	H04607	Hs.9218	ESTs			53.00		
	111747	AJ741471	Hs.23666	ESTs	46.20				
	111807	R33508	Hs.18827	ESTs	16.00				
40	111862	R37472	Hs.21559	EST				3.91	
	112045	AJ372588	Hs.8022	TU3A protein				2.74	
	112057	R43713	Hs.22945	EST				4.92	
	112214	AW148652	Hs.167398	ESTs			13.00		
	112263	R52393	Hs.25917	ESTs		2.43			
45	112314	AW206093	Hs.748	ESTs	9.00				
	112324	R55965	Hs.26479	limbic system-associated membrane protei			14.00		
	112362	AW300887	Hs.26638	ESTs; Weakly similar to CD20 receptor [H		2.49			
	112380	H63010	Hs.5740	ESTs		2.34			
	112425	AA324998	Hs.321677	ESTs; Weakly similar to III ALU SUBFAMI	8.00				
50	112473	R65993	Hs.279798	pregnancy specific beta-1-glycoprotein 9			29.80	4.53	
	112492	N51620	Hs.28694	ESTs				3.62	
	112541	AF038392	Hs.116674	ESTs					
	112620	R80552	Hs.29040	ESTs		2.37			
	112623	AW373104	Hs.25094	ESTs		2.26			
55	112867	T03254	Hs.167393	ESTs			12.00		
	112894	T08188	Hs.3770	ESTs	6.50				
	112954	AA928953	Hs.6655	ESTs	7.00				
	113029	AW081710	Hs.7369	ESTs; Weakly similar to III ALU SUBFAMI				4.39	
	113086	AA346839	Hs.209100	DKFZP434C171 protein				4.47	
60	113140	T50405	Hs.175967	ESTs			10.00		
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	14.00				
	113257	AJ821378	Hs.159367	ESTs				3.72	
	113394	T81473	Hs.177894	ESTs				3.60	
	113437	T85349	Hs.15923	EST	35.00				
65	113454	AJ022166	Hs.16188	ESTs	6.00				
	113502	T89130		ESTs	39.60				
	113552	AJ654223	Hs.16026	ESTs				3.88	
	113645	T95358	Hs.333181	ESTs				2.58	
	113691	T96935	Hs.17932	EST			38.20		
70	113706	AA004693	Hs.269192	ESTs				3.09	
	113883	U89281	Hs.11958	oxidative 3 alpha hydroxysteroid dehydro		2.31			
	113924	BE178285	Hs.170056	Homo sapiens mRNA; cDNA DKFZp586B0220 (f	30.40				
	114035	W92798	Hs.269181	ESTs			13.00		
	114058	AK002018	Hs.114727	ESTs				5.00	
75	114084	AA708035	Hs.12248	ESTs			40.60		
	114121	H05785	Hs.25425	ESTs		2.31			
	114124	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)	7.00				
	114275	AW515443	Hs.306117	interleukin 13 receptor; alpha 1	6.00				
	114297	AA149707	Hs.173091	DKFZP434K151 protein			48.80		
80	114427	AA017176	Hs.33532	ESTs; Highly similar to Miz-1 protein [H				3.45	
	114449	AA020736		*ze63b11.s1 Soares retina N2b4HR Homo sa			10.00		
	114452	AJ369275	Hs.243010	ESTs; Moderately similar to RTC0_HUMAN G	14.00				
	114609	AA079505		*zm97a5.s1 Stratagene colon HT29 (#93722				3.13	
	114648	AA101056		*zm25b3.s1 Stratagene neuroepithelium NT			35.40		
85	114731	BE094291	Hs.155651	Homo sapiens HNF-3beta mRNA for hepatocy				3.42	
	114762	AA146979	Hs.288464	ESTs	33.00				

	114776	AA151719	Hs.95834	ESTs	34.40				
	115009	AA251561	Hs.48689	ESTs	30.20				
	115272	AW015947		ESTs; Weakly similar to hypothetical L1	32.60				
5	115279	AW964897	Hs.290825	ESTs		6.00			
	115302	AL109719	Hs.47578	ESTs			12.00		
	115365	AW976252	Hs.258391	ESTs				3.32	
	115559	AL079707	Hs.207443	ESTs			48.00		
	115566	AI142336	Hs.43977	ESTs			56.20		
10	115683	AF255910	Hs.54650	ESTs; Weakly similar to (define not ava	31.40				
	115744	AA418538	Hs.43945	ESTs; Highly similar to dJ1178H5.3 [Hsa			33.60		
	115819	AA486620	Hs.41135	Endomucin 2			74.40		
	115949	AI478427	Hs.43125	ESTs		3.18			
	115965	AA001732	Hs.173233	ESTs			388.80		
15	116035	AA621405	Hs.184664	ESTs			33.20		
	116049	AA454033	Hs.41644	ESTs			45.80		
	116081	AI190071	Hs.55278	ESTs				3.57	
	116082	AB029496	Hs.59729	ESTs		3.06			
	116213	AA292105	Hs.326740	leucine rich repeat (in FLII) interactin	50.60				
20	116228	AI767947	Hs.50841	ESTs; Weakly similar to tuftelin [M.musc		3.85			
	116250	N76712	Hs.44829	ESTs		6.00			
	116419	AI613480	Hs.47152	ESTs; Weakly similar to testicular tekli			30.00		
	116617	D80761	Hs.45220	EST		2.27			
	116784	AB007979	Hs.301281	tenascin R (restrictin; janusin)	47.20				
25	116835	N39230	Hs.38218	ESTs			41.20		
	116970	AB023179	Hs.9059	KIAA0962 protein			91.00	11.00	
	117023	AW070211	Hs.102415	ESTs					
	117027	AW085208	Hs.130093	ESTs	49.40		32.60		
30	117036	H88908	Hs.41192	EST		8.67			
	117110	AA160079	Hs.172932	ESTs			30.60		
	117209	W03011	Hs.306881	ESTs				9.29	
	117325	N23599	Hs.43398	ESTs					3.19
	117454	N29569	Hs.44055	ESTs					
	117475	N30205	Hs.93740	ESTs	44.00				
35	117543	BE218453	Hs.42722	ESTs		16.00			
	117567	AW444761	Hs.44565	ESTs			12.00		
	117570	N48649	Hs.44583	ESTs			11.00		
	117600	N34963	Hs.44676	EST				3.74	
	117730	N45513	Hs.46608	ESTs		6.00			
40	117791	N48325	Hs.93956	EST		9.00			
	117929	N51075	Hs.47191	ESTs			29.20		
	117990	AA446167	Hs.47385	ESTs		8.00			
	118224	N62275	Hs.48503	EST	31.40				
45	118244	N62516	Hs.48556	ESTs	32.80				
	118357	AL109567	Hs.124154	Homo sapiens mRNA full length insert cDN		2.40			
	118446	N66361	Hs.269121	ESTs		2.28			
	118447	N66399	Hs.49193	EST	30.80				
	118530	N67900	Hs.118446	ESTs				3.10	
	118549	N68163	Hs.322954	EST				3.41	
50	118823	W03754	Hs.50813	ESTs; Weakly similar to long chain fatty		3.94			
	118862	W17065	Hs.54522	ESTs				3.58	
	118935	AI979247	Hs.247043	KIAA0525 protein			33.00		
	118944	AI734233	Hs.226142	ESTs; Weakly similar to !!!! ALU SUBFAM				11.43	
	118995	N94591	Hs.323056	ESTs		14.00			
55	119073	BE245360	Hs.279477	ERG-2/ERG-1; V-ets avian erythroblastosi			52.60		
	119268	T16335	Hs.65325	EST	31.40				
	119514	W37937		Accession not listed in Genbank				3.50	
	119824	W74536	Hs.184	advanced glycosylation end product-speci		2.75			
	119831	AL117694	Hs.58419	DKFZP586L2024 protein					3.21
60	119861	W78816	Hs.49943	ESTs; Moderately similar to !!!! ALU SUB			33.80		
	119889	W84346	Hs.58671	ESTs			30.03		
	119921	W86192	Hs.58815	ESTs	29.00				
	120082	H80286	Hs.40111	ESTs				3.80	
	120094	AA811339	Hs.124049	ESTs		6.00			
65	120132	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)			36.60		
	120378	AA223249	Hs.285728	ESTs		12.00			
	120404	AB023230	Hs.96427	KIAA1013 protein	39.40				
	120504	AA256837		ESTs			8.00		
70	120512	N55761	Hs.194718	ESTs	33.00				
	120667	AA287740	Hs.78335	microtubule-associated protein; RP/EB fa				4.18	
	120777	AA287702	Hs.10031	KIAA0955 protein			46.60		
	121082	AA398722		ESTs			39.00		
	121191	AA400205	Hs.104447	ESTs	41.60				
	121248	AA400914	Hs.97827	EST				5.08	
75	121363	AI287280	Hs.97933	ESTs			12.00		
	121366	AI743515		ESTs			20.00		
	121483	AI660332	Hs.25274	ESTs; Moderately similar to putative sev				3.32	
	121518	AA412155		ESTs			30.20		
	121545	AA412442	Hs.98132	ESTs		2.29			
80	121622	AA416931	Hs.126065	ESTs		9.00			
	121665	AA416556	Hs.98234	ESTs			34.80		
	121709	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	34.80				
	121730	AI140683	Hs.98328	ESTs					
	121740	AA421138	Hs.98334	EST		7.00			
85	121772	AI590770	Hs.110347	Homo sapiens mRNA for alpha integrin bin	36.20				
	121821	AL040235	Hs.3346	ESTs				3.61	

WO 02/086443

PCT/US02/12476

	121835	AB033030	Hs.300670	ESTs	2.34				
	121841	AA427794	Hs.104864	ESTs	2.61				
	121885	AA934883	Hs.98467	ESTs				2.25	
	121888	AA426429	Hs.98463	ESTs				2.92	
5	121938	AA428659	Hs.98610	ESTs			46.80		
	121950	AA429515		EST			31.40		
	122030	AA431310	Hs.98724	ESTs	34.40				
	122054	AA431725	Hs.98746	EST				3.58	
10	122211	AA300900	Hs.98849	ESTs; Moderately similar to bithoraxoid-	49.40				
	122233	AA436455	Hs.98872	EST	29.80				
	122247	AA436676	Hs.98890	EST			39.80		
	122253	AA436703	Hs.104936	ESTs; Weakly similar to hypothetical pro	9.00				
	122266	AA436840	Hs.98907	EST				3.60	
	122285	AA436981	Hs.121602	EST				3.14	
15	122409	AA446830	Hs.99081	ESTs	30.80				
	122485	AA524547	Hs.160318	phospholemman		2.65			
	122697	AA420683	Hs.98321	Homo sapiens cDNA FLJ14103 fis, clone MA	15.00				
	122772	AW117452	Hs.99489	ESTs	6.57				
20	122831	AJ857570	Hs.5120	ESTs				3.37	
	122913	AJ638774	Hs.105328	ESTs			32.20		
	123049	BE047680	Hs.211869	ESTs			41.80		
	123076	AJ345569	Hs.190046	ESTs	35.80				
	123136	AW451999	Hs.194024	ESTs				2.58	
	123309	N52937	Hs.102679	ESTs					
25	123455	AA353113	Hs.112497	ESTs			82.80		
	123691	AA609579	Hs.112724	ESTs				3.95	
	123756	AA609971	Hs.112795	EST	35.40				
	123802	AA620448		Homo sapiens clone 24760 mRNA sequence	58.00				
30	123837	AJ807243	Hs.112893	ESTs			32.40		
	123844	AA938905	Hs.120017	olfactory receptor, family 7; subfamily		2.63			
	123936	NM_004673	Hs.241519	ESTs	29.00				
	123987	C21171	Hs.95497	ESTs; Weakly similar to GLUCOSE TRANSPOR			70.60		
	124013	AJ521936	Hs.107149	ESTs; Weakly similar to PTB-ASSOCIATED S	28.40				
35	124160	R40290	Hs.124685	ESTs			13.00		
	124205	H77570	Hs.108135	ESTs				4.74	
	124226	AA618527	Hs.190266	ESTs		2.35			
	124246	H67680	Hs.270962	ESTs			29.40		
	124348	AJ796320	Hs.10299	ESTs	17.00				
40	124358	AW070211	Hs.102415	gamma35g11.s1 Morton Fetal Cochlea Homo sa		3.07			
	124409	AJ814166	Hs.107197	ESTs				3.14	
	124442	AW663632	Hs.285625	TATA box binding protein (TBP)-associate		2.48			
	124468	N51413	Hs.109284	ESTs			30.80		
	124479	AB011130	Hs.127436	calcium channel; voltage-dependent; alph		2.50			6.03
45	124519	AJ670056	Hs.137274	ESTs; Weakly similar to SPLICEOSOME ASSO					
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	59.20				
	124866	AJ768289	Hs.304389	ESTs		8.00			
	124874	BE550182	Hs.127826	ESTs			37.60		
	125097	AW576389	Hs.335774	ESTs			10.00		
50	125179	AW206468	Hs.103118	ESTs				3.12	
	125200	AW836591	Hs.103156	ESTs					2.79
	125299	T32982	Hs.102720	ESTs			34.20		
	125400	AL110151	Hs.128797	DKFZP586D0824 protein	29.00				
	125810	H00083		aryl hydrocarbon receptor-interacting pr	32.20				
55	126176	BE242256	Hs.2441	KIAA0022 gene product		12.00			
	126303	D78841		HUM525A059 Human placenta polyA+ (TFu)			33.60		
	126403	AW629054	Hs.125976	ESTs; Weakly similar to metalloprotease/	35.80				
	126507	AL040137	Hs.23964	ESTs; Weakly similar to HC1 ORF (Mmuscu			29.80		
	126773	AA648284	Hs.187584	ESTs					
60	127307	AW962712	Hs.126712	ESTs; Weakly similar to pL2 hypotheticala	39.60				
	127462	AA760778	Hs.293977	aa59b04.s1 NCL_CGAP_GCB1 Homo sapiens c	28.80		34.40		
	127486	AW002846	Hs.105468	ESTs		9.00			
	127572	AA594027	Hs.191788	ESTs		2.36			
	127609	X80031	Hs.530	ESTs			29.40		
	127832	AW976035	Hs.292396	ESTs			37.20		
65	127898	AA774725	Hs.128970	ESTs				4.42	
	128073	AW340720	Hs.125983	ESTs			38.40		
	128101	AA905730	Hs.128254	ESTs		7.33			
	128149	NM_012214	Hs.177576	mannosyl (alpha-1,3)-glycoprotein beta-				2.58	
70	128212	W27411	Hs.336920	glutathione peroxidase 3 (plasma)		3.09			
	128333	W68800	Hs.12126	ESTs; Weakly similar to LR8 (H.sapiens)			34.40		
	128364	N78462	Hs.269152	ESTs; Weakly similar to ZINC FINGER PROT		10.00			
	128426	AJ265784	Hs.145197	ESTs				4.31	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel; s	31.20				
75	128634	AA454918		ESTs; Moderately similar to !!!! ALU SUB			41.60		
	128687	AW271273	Hs.23767	ESTs			87.00		
	128726	AJ311238	Hs.104476	ESTs					
	128773	NM_004131	Hs.1051	granzyme B (granzyme 2; cytotoxic T-lymp			9.00		4.02
	128833	W26667	Hs.184581	ESTs					3.76
80	128870	H39537	Hs.75309	eukaryotic translation elongation factor		2.66			
	128878	R25513	Hs.10583	ESTs				3.10	
	128885	AF134803	Hs.180141	cofilin 2 (muscle)			11.00		
	128998	W04245	Hs.107761	ESTs; Weakly similar to PUTATIVE RHO/RAC				3.21	
	129000	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII inh					3.68
85	129038	AW156903	Hs.108124	ribosomal protein L41				3.17	
	129098	AW580945	Hs.330466	ESTs	34.60				

	129210	AL039940	Hs.202949	KIAA1102 protein				4.09
	129240	AA361258	Hs.237868	interleukin 7 receptor	2.29			
	129262	BE222198	Hs.109843	ESTs		3.30		
5	129301	AF182277	Hs.330780	Human cytochrome P450-11B (h11B3) mRNA;				4.05
	129331	AW167668	Hs.279772	ESTs; Highly similar to CGI-38 protein [4.09
	129381	AW245805	Hs.110903	claudin 5 (transmembrane protein deleted	2.93	160.80		
	129585	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		10.00		
	129595	U09550	Hs.1154	oviductal glycoprotein 1; 120kD				
10	129613	AW978517	Hs.172847	ESTs; Weakly similar to collagen alpha 1		3.40		
	129782	AW016932	Hs.104105	EST	9.00			
	129950	F07783	Hs.1369	decay accelerating factor for complement		87.80		
	129958	R27496	Hs.1378	annexin A3		44.60		
	129959	AL036554	Hs.274463	defensin; alpha 1; myeloid-related seque	2.72			
15	130160	AA305688	Hs.267695	UDP-GalbetaGlcNAc beta 1,3-galactosyltr		42.20		
	130259	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator	2.54			
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic; Dr		51.60		
	130312	AF056195	Hs.15430	DKFZP586G1219 protein			3.16	
	130436	NM_001928	Hs.155597	D component of complement (adipsin)				4.11
20	130523	AA999702	Hs.214507	ESTs			4.77	
	130799	AB028945	Hs.12696	ESTs	6.00			
	130885	NM_005883	Hs.20912	adenomatous polyposis coli like			3.54	
	131002	AL050295	Hs.22039	KIAA0758 protein				3.50
	131012	AL039940	Hs.202949	KIAA1102 protein	20.00			
25	131031	NM_001650	Hs.288650	aquaporin 4	41.20			
	131061	N64328	Hs.268744	ESTs; Moderately similar to KIAA0273 [H.		31.40		
	131066	AW169287	Hs.22588	ESTs		29.60		
	131082	AI091121	Hs.246218	ESTs; Weakly similar to zinc finger prot		9.00		
	131087	AF147709	Hs.22824	ESTs; Weakly similar to p160 myb-binding				3.86
30	131161	AF033382	Hs.23735	potassium voltage-gated channel; subfam			3.14	
	131179	AA171388	Hs.184482	DKFZP586D0624 protein			3.80	
	131182	AI824144	Hs.23912	ESTs				3.67
	131205	NM_003102	Hs.2420	superoxide dismutase 3; extracellular	2.98			
35	131277	AA131466	Hs.23767	ESTs	3.15			
	131281	AA251716	Hs.25227	ESTs		32.20		
	131282	X03350	Hs.4	alcohol dehydrogenase 3 (class I); gamma				3.44
	131285	AI567943	Hs.25274	ESTs; Moderately similar to putative sev			6.40	
	131355	R52804	Hs.25956	DKFZP564D206 protein	8.00			
40	131391	AW085781	Hs.26270	ESTs	10.00			
	131461	AA992841	Hs.27263	butyrate response factor 2 (EGF-response	28.80			
	131487	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f				4.03
	131517	AB037789	Hs.263395	ESTs; Highly similar to semaphorin VIa [39.00			
	131545	AL137432	Hs.28564	ESTs		11.00		
	131583	AK000383	Hs.323092	ESTs; Weakly similar to dual specificity		10.00		
45	131647	AA359615	Hs.30089	ESTs	2.47			
	131675	H15205	Hs.30509	ESTs			3.06	
	131676	AI126821	Hs.30514	ESTs	45.80			
	131708	S60415	Hs.30941	calcium channel; voltage-dependent; beta	2.28			
	131717	X94630	Hs.3107	CD97 antigen				3.78
50	131756	AA443966	Hs.31595	ESTs		40.60		
	131762	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII Inhi				3.67
	131821	AA017247	Hs.164577	ESTs	2.87			
	131839	AB014533	Hs.33010	KIAA0633 protein				3.48
	131861	AL088858	Hs.184245	KIAA0929 protein Msx2 Interacting nuclea	54.00			
55	132015	AI418006	Hs.3731	ESTs		49.20		
	132070	BE622641	Hs.38489	ESTs		34.80		
	132242	AA332697	Hs.42721	ESTs	2.68			
	132334	AW080704	Hs.45033	lacrimal proline rich protein	4.66			
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	34.20			
60	132490	NM_001290	Hs.4980	LIM binding domain 2		2.66		
	132533	AI922988	Hs.172510	ESTs	13.00			
	132598	X80031	Hs.530	collagen; type IV; alpha 3 (Goodpasture		30.60		
	132619	H28855	Hs.53447	ESTs; Moderately similar to kinesin ligh			4.02	
	132652	N41739	Hs.61260	ESTs			3.18	
65	132726	N52298	Hs.55608	ESTs; Weakly similar to cDNA EST yk484g1		11.43		
	133028	R51604	Hs.300842	ESTs	2.37			
	133071	BE384932	Hs.64313	ESTs	2.27			
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein	2.63			
	133129	AA428580	Hs.65551	ESTs				5.49
70	133147	AA026533	Hs.66	Interleukin 1 receptor-like 1	6.20			
	133151	NM_014051	Hs.94896	ESTs			3.69	
	133213	AA903424	Hs.6786	ESTs		31.40		
	133276	AW978439	Hs.69504	ESTs		9.00		
	133377	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene famil	41.20			
75	133407	AF017987	Hs.7306	secreted frizzled-related protein 1	50.20			
	133535	AL134030	Hs.284180	protocadherin 2 (cadherin-like 2)			3.72	
	133537	U41518	Hs.74602	aquaporin 1 (channel-forming integral pr				3.35
	133656	BE149455	Hs.75415	Accession not listed in Genbank	2.65			
	133689	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)		90.80		
80	133779	T58486	Hs.222566	ESTs			3.05	
	133978	AF035718	Hs.78061	transcription factor 21	2.92			
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec				3.45
	134000	AW175787	Hs.334841	selenium binding protein 1				4.05
	134111	AI372588	Hs.8022	TU3A protein	4.49			
85	134185	AA285136	Hs.301914	Homo sapiens mRNA; cDNA DKFZp586K1220 (f			3.27	
	134204	AI873257	Hs.7994	ESTs; Weakly similar to CGI-69 protein [40.80		

PCT/US02/12476

25 TABLE 2B shows the accession numbers for those primekeys lacking unigenes/D's for Table 2A. For each probe set we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

30 Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

	Pkey	CAT number	Accessions
35	108447	43452_-7	AA079126
	108550	120073_1	AA084867 AA084996
	108655	127522_1	AA099960 AA113013
	102397	44371_-1	U41898
	126303	1525933_1	D78841 D78880
40	125810	1554054_1	H00083 R81062
	103627	2615_2	Z48513 Z48512
	121366	280401_1	A1743515 AA405617 AW276705
	114609	116777_1	AA079505 AA079537
	115272	172113_1	AW015947 AA211890 AA279425
	108338	112186_1	AA070773 AA070774
45	108434	114012_1	AA078899 AA078782 AA075788
	123802	genbank_AA620448	AA620448
	102310	NOT_FOUND_entrez_U33839	U33839
	102636	entrez_U67092 U67092	
	104776	genbank_AA026349	AA026349
50	120504	genbank_AA256837	AA256837
	113502	genbank_T89130T89130	
	108499	genbank_AA083103	AA083103
	101308	entrez_L41390 L41390	
55	108529	genbank_AA102425	AA102425
	103098	Z21_215	M86361 Z26593 X02850 D13070 AE000659 M17649 M87869 M87871 X61077 M16286 AF018169 X61079 S59351 X60142 AF043169
	103241	entrez_X76223 X76223	
	103508	entrez_Y10141 Y10141	
	103575	entrez_Z26256 Z26256	
60	119514	NOT_FOUND_entrez_W37937 W37937	
	121082	genbank_AA398722	AA398722
	128634	AA464918_at AA464918	
	105817	genbank_AA397825	AA397825
	121518	genbank_AA412155	AA412155
	114449	genbank_AA020736	AA020736
65	114648	genbank_AA101056	AA101056
	121950	genbank_AA429515	AA429515
	107723	genbank_AA015967	AA015967

Table 3A shows 452 genes up-regulated in chronically diseased lung relative to normal lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number					
	ExAccn:	Exemplar Accession number, Genbank accession number					
	UnigenelD:	Unigene number					
	Unigene Title:	Unigene gene title					
10	R1:	80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of AI for normal lung samples.					
	R2:	80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas					
	R3:	70th percentile of AI for chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples					
15							
	Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2	R3
20	135423	U50531	Hs.138751	Human BRCA2 region, mRNA sequence CG030	12.40		
	135378	AW961818	Hs.24379	MUM2 protein			2.13
	135346	NM_000928	Hs.992	phospholipase A2, group IB (pancreas)			
	135235	AW298244	Hs.293507	ESTs	12.40		
	135057	U90268	Hs.93810	cerebral cavernous malformations 1	11.67		
25	134951	BE305081	Hs.169358	hypothetical protein		8.00	
	134799	M36821	Hs.89690	GRO3 oncogene		8.20	
	134786	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous			
	134772	NM_000829	Hs.163697	glutamate receptor, ionotropic, AMPA 4	29.80		
	134752	BE246762	Hs.89499	arachidonate 5-lipoxygenase			1.93
	134749	T28499	Hs.89485	carbonic anhydrase IV			2.07
30	134696	BE326276	Hs.8861	ESTs			
	134636	NM_005582	Hs.87205	lymphocyte antigen 64 (mouse) homolog, r	13.60		
	134627	AI018768	Hs.12482	glyceronephosphate O-acyltransferase			1.92
	134622	AW975159	Hs.293097	ESTs, Weakly similar to A55380 facitogeni			1.92
35	134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	13.20		
	134561	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h			1.78
	134468	NM_001772	Hs.83731	CD33 antigen (gp67)		6.20	
	134417	NM_006416	Hs.82921	solute carrier family 35 (CMP-sialic aci			
	134343	D50683	Hs.82028	transforming growth factor, beta recepto			
40	134323	BE170651	Hs.8700	deleted in liver cancer 1			
	134300	NM_001430	Hs.8136	endothelial PAS domain protein 1			
	134299	AW580939	Hs.97199	complement component C1q receptor			
	134253	X52075	Hs.80738	sialophorin (gpL115, leukosialin, CD43)	20.60		
	134182	D52059	Hs.7972	KIAA0871 protein	12.20		
45	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec			
	133978	AF035718	Hs.78061	transcription factor 21			
	133835	AI677897	Hs.76640	RGC32 protein			
	133651	AI301740	Hs.173381	dihydropyrimidinase-like 2			
	133633	D21262	Hs.75337	nucleolar and colloid-body phosphoprotein	15.20		
50	133565	AW955776	Hs.313500	ESTs, Moderately similar to ALU7_HUMAN A			
	133548	AW946384	Hs.178112	DNA segment, single copy probe LNS-CA/L			1.77
	133488	AA335295	Hs.74120	adipose specific 2			
	133478	X83703	Hs.31432	cardiac ankyrin repeat protein			2.08
	133337	AF085983	Hs.293676	ESTs		9.60	
55	133200	AB037715	Hs.183639	hypothetical protein FLJ10210			1.77
	133153	AF070592	Hs.66170	HSKM-B protein	30.60		
	133130	AI128606	Hs.6557	zinc finger protein 161	22.60		
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein			
	132928	AW168082	Hs.169449	protein kinase C, alpha	13.80		
60	132836	AB023177	Hs.29900	KIAA0960 protein			
	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,	41.60		
	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypothei	40.40		
	132548	X12830	Hs.193400	interleukin 6 receptor		7.20	
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi		4.76	
	132439	AK001942	Hs.4863	hypothetical protein DKFZp566A1524			1.88
65	132240	AB018324	Hs.42676	KIAA0781 protein	21.20		
	132210	NM_007203	Hs.42322	A kinase (PRKA) anchor protein 2			1.99
	132199	AL041299	Hs.165084	ESTs	15.20		
	131751	T96555	Hs.31562	ESTs			1.76
70	131745	AI828559	Hs.31447	ESTs, Moderately similar to A46010 X-II	27.80		
	131694	NM_000246	Hs.3076	MHC class II transactivator		4.00	
	131686	NM_012296	Hs.30687	GRB2-associated binding protein 2			
	131676	AI126821	Hs.30514	ESTs		6.20	
	131629	Z45794	Hs.238809	ESTs	21.40		
75	131589	C18825	Hs.29191	epithelial membrane protein 2			
	131536	AA019201	Hs.269210	ESTs		9.40	
	131517	AB037789	Hs.263395	sema domain, transmembrane domain (TM),		3.59	
	131355	RS2804	Hs.25956	DKFZP564D206 protein		4.48	
	131253	R71802	Hs.24853	ESTs	15.00		
80	131207	AF104266	Hs.24212	latrophilin			1.75
	131156	AI472209	Hs.323117	ESTs			1.84
	131066	AW169287	Hs.22588	ESTs		3.54	
	131061	N64328	Hs.268744	KIAA1796 protein			
	131053	AA348541	Hs.296261	guanine nucleotide binding protein (G pr			1.93
85	130895	AA641767	Hs.21015	hypothetical protein DKFZp564L0864 simil	16.60		
	130762	D84371	Hs.1898	paraonoxase 1	12.00		

	130657	AW337575	Hs.201591	ESTs		
	130655	AI831952	Hs.17409	cysteine-rich protein 1 (intestinal)		
	130589	AL110226	Hs.16441	DKFZP434H204 protein		2.08
5	130562	D50402	Hs.182611	solute carrier family 11 (proton-coupled)		1.91
	130555	R69743	Hs.116774	Integrin, alpha 1	9.60	
	130365	W56119	Hs.155103	eukaryotic translation initiation factor	11.60	
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic, Dr	6.60	
	130259	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator		1.91
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 18)	21.20	
	129958	R27496	Hs.1378	annexin A3	5.05	
	129898	AI672731	Hs.13256	ESTs		
	129875	AA181018	Hs.13056	hypothetical protein FLJ13920	18.60	
	129699	AB007899	Hs.12017	homolog of yeast ubiquitin-protein ligase		
15	129626	F13272	Hs.111334	ferritin, light polypeptide		
	129598	N30435	Hs.11556	Homo sapiens cDNA FLJ12556 fis, clone NT	22.63	
	129593	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f		
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		2.53
	129527	AA769221	Hs.270847	delta-tubulin	39.20	
20	129402	W72062	Hs.11112	ESTs		2.11
	129385	AA172106	Hs.110950	Rag C protein	15.20	
	129315	NM_014563	Hs.174038	spondyloepiphyseal dysplasia, late	12.40	
	129312	T97579	Hs.110334	ESTs, Weakly similar to I78885 serine/th	20.83	
	129240	AA361258	Hs.237868	Interleukin 7 receptor		1.95
25	129210	AL039940	Hs.202949	KIAA1102 protein		
	129122	AW958473	Hs.301957	nudix (nucleoside diphosphate linked moi	4.20	
	129057	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)		
	128946	Y13153	Hs.107318	kynurenine 3-monooxygenase (kynurenine 3	5.20	
	128798	AF015525	Hs.302043	chemokine (C-C motif) receptor-like 2		
30	128789	AW368576	Hs.139851	caveolin 2		2.24
	128778	AA504776	Hs.186709	ESTs, Weakly similar to I38022 hypothet	12.20	
	128766	AW160432	Hs.296460	craniofacial development protein 1	26.40	
	128631	R44239	Hs.155546	KIAA1080 protein; Golgi-associated, gamma		1.78
	128624	BE154765	Hs.102647	ESTs, Weakly similar to TRHY_HUMAN TRICH		2.51
35	128609	NM_003616	Hs.102456	survival of motor neuron protein interac	16.00	
	128603	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	12.80	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel, s	4.00	
	128458	H55864	Hs.56340	ESTs		
40	128051	AF150882	Hs.186877	sodium channel, voltage-gated, type XII,	17.20	
	127968	AA830201	Hs.124347	ESTs	21.30	
	127959	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L		
	127944	AI557081	Hs.262476	S-adenosylmethionine decarboxylase 1	10.60	
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	13.40	
	127896	AI669586	Hs.222194	ESTs	7.00	
45	127859	AA761802	Hs.291559	ESTs	14.00	
	127817	AA836641	Hs.163085	ESTs	14.00	
	127742	AW293496	Hs.180138	ESTs	11.00	
	127628	AI240102	Hs.322430	NDRG family, member 4	11.10	
50	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture		
	127582	AA908954	Hs.130944	ESTs	19.60	
	127543	AK000787	Hs.157392	Homo sapiens cDNA FLJ20780 fis, clone CO	15.40	
	127535	AA568424	Hs.164450	ESTs	17.50	
	127404	AI379920	Hs.270224	ESTs	14.60	
55	127396	L31968	Hs.187991	DKFZP564A122 protein	15.40	
	127374	AA442797	Hs.312110	ESTs, Weakly similar to I38022 hypothet	14.60	
	127346	AA203616	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe	21.00	
	127340	BE047653	Hs.119183	ESTs, Weakly similar to ZN91_HUMAN ZINC	15.80	
	127307	AW962712	Hs.126712	ESTs, Weakly similar to AF191020 1 E2IG5		
60	127242	AW390395	Hs.181301	cathepsin S	22.60	
	127167	AA625690	Hs.190272	ESTs	21.40	
	127046	AA321948	Hs.293968	ESTs	41.20	
	126928	AA480902	Hs.137401	ESTs	11.00	
	126900	AF137386	Hs.12701	plasmolipin		1.78
65	126852	AA399961		gbzu68c01.r1 Soares_testis_NHT Homo sap	5.60	
	126816	AA248234		gbzsg2228.seq.F Human fetal heart, Lamb	12.20	
	126812	AB037860	Hs.173933	nuclear factor I/A	17.19	
	126666	AA648886	Hs.151999	ESTs	13.57	
	126645	AA316181	Hs.61635	stx transmembrane epithelial antigen of	15.40	
70	126592	AI611153	Hs.6093	Homo sapiens cDNA: FLJ22783 fis, clone K	4.67	
	126556	AF255303	Hs.112227	membrane-associated nucleic acid binding	18.00	
	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	16.77	
	126299	AW979155	Hs.298275	amino acid transporter 2	14.60	
	126218	AL049801	Hs.13649	Novel human gene mapping to chromosome 13	3.50	
75	126182	AA721331	Hs.293771	ESTs	13.40	
	126177	AW752782	Hs.129750	hypothetical protein FLJ10546	18.20	
	126142	H86261	Hs.40568	ESTs	14.00	
	126077	M78772	Hs.210836	ESTs	16.59	
	125994	AI990529	Hs.270799	ESTs	17.40	
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	13.00	
80	125847	AW161885	Hs.249034	ESTs	49.57	
	125831	HD04043		gb:bj45c03.r1 Soares placenta Nb2HP Homo		
	125731	R61771	Hs.26912	ESTs	13.20	
	125676	BE612918	Hs.151973	hypothetical protein FLJ23511	11.20	
	125661	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S		
85	125552	H09701	Hs.278366	ESTs, Weakly similar to I38022 hypotheti	12.60	
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A	33.40	

	125422	AA903229	Hs.153717	ESTs		1.80
	125331	AI422396	Hs.161378	ESTs	38.00	
	125309	T12411	Hs.183745	hypothetical protein FLJ13456	18.20	
5	125167	AL137540	Hs.102541	netrin 4		1.95
	125139	AW194933	Hs.9788	hypothetical protein MGC10924 similar to		1.84
	125042	T78906	Hs.269432	ESTs, Moderately similar to ALU1_HUMAN	21.80	
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	10.60	
	124631	NM_014053	Hs.270594	FLVCR protein	23.20	
10	124578	N58321	Hs.231500	EST	21.43	
	124574	AL036596	Hs.42322	A kinase (PRKA) anchor protein 2		1.77
	124472	N52517	Hs.102670	EST	37.20	
	124438	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A		
	124357	N22401		gb:vw37g07.s1 Morton Fetal Cochlea Homo	14.64	
15	124306	AW973078	Hs.293039	ESTs		4.00
	124214	H58608	Hs.151323	ESTs		
	124097	AW298235	Hs.101689	ESTs	27.20	
	123978	T89832	Hs.170278	ESTs		2.03
	123972	T46848	Hs.70337	immunoglobulin superfamily, member 4	6.00	
20	123961	AL050184	Hs.21610	DKFZP434B203 protein		1.79
	123936	NM_004673	Hs.241519	angiotensin-like 1	15.80	
	123802	AA620448		gb:ae58c09.s1 Stratagene lung carcinoma	4.23	
	123734	AA609861	Hs.312447	ESTs	4.20	
	123619	AA602984		gb:nc97c02.s1 NCI_CGAP_Pr2 Homo sapiens	33.60	
25	123596	AA421130	Hs.112640	EST	10.93	
	123476	AA384564	Hs.108829	ESTs		2.18
	123340	AA504264	Hs.182937	peptidylprolyl isomerase A (cyclophilin	11.20	
	123190	AA489212	Hs.105228	EST	14.20	
	123136	AW451999	Hs.194024	ESTs		7.00
30	123073	AA485061	Hs.105652	ESTs	31.20	
	123055	AA482005	Hs.105102	ESTs, Weakly similar to reverse transcri		4.80
	122699	AA456130	Hs.301721	KIAA1255 protein		5.00
	122679	AA811286	Hs.192837	ESTs, Weakly similar to ALU5_HUMAN ALU S	14.40	
	122633	NM_001546	Hs.34853	Inhibitor of DNA binding 4, dominant neg		
35	122553	AA451884	Hs.190121	ESTs	40.00	
	122544	AW973253	Hs.292689	ESTs	15.40	
	122485	AA524547	Hs.160318	FXYD domain-containing ion transport reg		1.81
	122211	AA300900	Hs.98849	ESTs, Moderately similar to AF161511.1 H	12.10	
	122127	AW207175	Hs.106771	ESTs		1.95
40	122011	AA431082		gb:zw78a10.s1 Soares_testis_NHT Homo sap		1.89
	121992	AI860775	Hs.98506	ESTs	3.60	
	121989	W58487	Hs.193784	Homo sapiens mRNA; cDNA DKFZp586K1922 (f		2.01
	121835	AB033030	Hs.300670	KIAA1204 protein		1.85
	121726	AF241254	Hs.178098	angiotensin I converting enzyme (peptidy	12.43	
45	121690	AV660305	Hs.110288	ESTs		1.82
	121643	AA640987	Hs.193767	ESTs		
	121633	AA417011	Hs.98175	EST	14.00	
	121622	AA416931	Hs.126065	ESTs	16.40	
	121497	AA412031	Hs.97901	EST	11.20	
50	121351	AW206227	Hs.287727	hypothetical protein FLJ23132	12.20	
	121314	W07343	Hs.182538	phospholipid scramblase 4		1.83
	121242	AA400857	Hs.97509	ESTs	22.40	
	121059	AA393283		gb:z74e03.r1 Soares_testis_NHT Homo sap	14.80	
55	120934	AA226198		gb:nc26a07.s1 NCI_CGAP_Pr1 Homo sapiens	21.20	
	120755	AA312934	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone		1.79
	120637	AA811804		gb:cb39a05.s1 NCI_CGAP_GC81 Homo sapiens	20.00	
	120484	AA253170	Hs.96473	EST	40.20	
	120336	N85785	Hs.181165	eukaryotic translation elongation factor		6.60
60	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	16.80	
	120132	W57554	Hs.125019	ESTs		4.73
	120041	AA830882	Hs.59368	ESTs		1.75
	119996	W88996	Hs.59134	EST	7.20	
	119970	AA767718	Hs.93581	hypothetical protein FLJ10512	11.20	
65	119861	W78816	Hs.49943	ESTs, Weakly similar to S65657 alpha-1C-	3.78	
	119824	W74536	Hs.184	advanced glycosylation end product-speci		
	119740	AW021407	Hs.21068	hypothetical protein	20.20	
	119271	AI061118	Hs.65328	Fanconi anemia, complementation group F	15.20	
	119221	C14322	Hs.250700	trypsin beta 1		
	119126	R45175	Hs.117183	ESTs	12.60	
70	119073	BE245360	Hs.279477	ESTs		
	118928	AA312799	Hs.283689	activator of CREM in testis	10.00	
	118901	AW292577	Hs.94445	ESTs	3.96	
	118661	AL137554	Hs.49927	protein kinase NYD-SP15	9.60	
	118607	AI377444	Hs.54245	ESTs, Weakly similar to S65824 reverse t	10.40	
75	118449	AI813865	Hs.164478	hypothetical protein FLJ21939 similar to		1.90
	118416	N66028	Hs.49105	FKBP-associated protein	16.20	
	118379	N54491	Hs.48990	ESTs		4.00
	118329	N63520		gb:yy62f01.s1 Soares_multiple_sclerosis_	6.60	
	118320	N63451	Hs.141600	ESTs, Weakly similar to alternatively s	3.80	
80	118253	AA497044	Hs.20887	hypothetical protein FLJ10392	17.60	
	118124	N56968	Hs.46707	chromosome 21 open reading frame 37	14.00	
	118056	AB037746	Hs.42768	hypothetical protein DKFZp76100113		1.86
	118032	N52802	Hs.47544	EST	5.00	
	117840	T26379	Hs.48802	Homo sapiens clone 23632 mRNA sequence	4.00	
85	117404	N39725	Hs.15220	zinc finger protein 106		1.90
	117314	N32498	Hs.42829	ESTs	14.20	

	117209	W03011	Hs.306881	MSTP043 protein		
	117023	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	2.31	
	116814	H50834		gb:yp86a10.s1 Soares fetal liver spleen	20.20	
5	116784	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	3.51	
	116766	AI608657	Hs.95097	ESTs	16.20	
	116712	AW901618	Hs.61935	Homo sapiens mRNA; cDNA DKFZp761071 (fr	6.80	
	116707	H10344	Hs.49050	ESTs, Weakly similar to A Chain A, Human	18.60	
	116351	AL133623	Hs.82501	similar to mouse Xrn1 / Dhm2 protein	19.40	
10	116279	AW971248	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S		
	116166	AL039940	Hs.202949	KIAA1102 protein	2.13	
	116152	AL040521	Hs.15220	zinc finger protein 106	1.75	
	116117	BE613410	Hs.31575	SEC63, endoplasmic reticulum translocon	13.20	
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	30.11	
15	115965	AA001732	Hs.173233	hypothetical protein FLJ10970	2.36	
	115955	AF263613	Hs.44198	Intracellular membrane-associated calciu	18.20	
	115844	AI373062	Hs.332938	hypothetical protein MGC5370	18.57	
	115883	AF255910	Hs.54650	junctional adhesion molecule 2	23.00	
	115673	AA406341	Hs.269908	Homo sapiens cDNA FLJ11991 fis, clone HE	11.82	
20	115672	AI889110	Hs.73251	ESTs	10.60	
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1	1.76	
	115313	AA080001	Hs.184411	albumin	25.20	
	115279	AW964897	Hs.290825	ESTs	8.00	
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	1.80	
25	115110	AK001671	Hs.11387	KIAA1453 protein	14.20	
	114999	BE246481	Hs.87856	ESTs	19.20	
	114930	AA237022	Hs.188717	ESTs	5.60	
	114922	AA235672	Hs.87491	ESTs	3.60	
	114837	BE244930	Hs.166895	ESTs	43.70	
30	114769	AA149060	Hs.296100	ESTs	11.00	
	114761	AA143781	Hs.126280	hypothetical protein FLJ23393	14.00	
	114736	AI610347	Hs.103812	ESTs, Moderately similar to ALU1_HUMAN A	4.20	
	114596	AA310162	Hs.169248	cytochrome c	10.71	
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	20.40	
35	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.40	
	114452	AI359275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE	17.20	
	114359	NM_016929	Hs.283021	chloride intracellular channel 5	2.09	
	114357	R41677	Hs.6107	Homo sapiens cDNA FLJ14839 fis, clone OV	12.40	
40	114251	H15261	Hs.21948	ESTs	2.00	
	114138	AW384793	Hs.15740	Homo sapiens mRNA; cDNA DKFZp434E033 (fr	11.40	
	114124	W57554	Hs.125019	ESTs	6.04	
	113946	AW083883	Hs.37896	Homo sapiens cDNA FLJ13510 fis, clone PL	1.82	
	113695	T96965	Hs.17948	ESTs, Weakly similar to ALUB_HUMAN !!!!		
45	113606	NM_013343	Hs.278951	NAG-7 protein	2.15	
	113590	R49642	Hs.142447	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.60	
	113560	T91015	Hs.268626	ESTs	32.00	
	113552	AI654223	Hs.16026	hypothetical protein FLJ23191		
	113540	AW152618	Hs.16757	ESTs		
50	113502	T89130		gb:ye12d01.s1 Stratagene lung (937210) H	8.35	
	113288	AI076838	Hs.12967	ESTs	12.40	
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	4.27	
	113238	R45467	Hs.189813	ESTs		
	113203	AA743563	Hs.10305	ESTs	21.20	
	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	1.92	
55	113089	T40707	Hs.270862	ESTs	14.33	
	113076	AF033199	Hs.8198	zinc finger protein 204	6.00	
	113009	T23699	Hs.7246	ESTs	9.40	
	112937	AI694320	Hs.6295	ESTs, Weakly similar to T17248 hypotheti	12.20	
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-f		
60	112794	R97018		gb:yy74b08.s1 Soares fetal liver spleen	26.60	
	112691	R88708	Hs.220647	ESTs	15.33	
	112602	AW004045	Hs.203365	ESTs	15.60	
	112366	AF035318	Hs.12533	Homo sapiens clone 23705 mRNA sequence	15.40	
	112210	R49645	Hs.7004	ESTs	14.00	
65	112064	AL049390	Hs.22689	Homo sapiens mRNA; cDNA DKFZp586O1318 (f	13.00	
	111998	R42379	Hs.138283	ESTs	11.00	
	111987	NM_015310	Hs.6763	KIAA0942 protein	22.40	
	111803	AA593731	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.77	
	111737	H04607	Hs.9218	ESTs	1.86	
70	111605	T91081	Hs.194178	ESTs, Moderately similar to PC4259 ferri	23.00	
	111510	R07856	Hs.16355	ESTs	11.02	
	111341	AL157484	Hs.22483	Homo sapiens mRNA; cDNA DKFZp762M127 (fr	1.88	
	111280	AA373527	Hs.19385	CGI-58 protein	18.40	
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f		
75	111232	AI247763	Hs.16928	ESTs	27.60	
	110942	R63503	Hs.28419	ESTs	14.80	
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	24.71	
	110837	H03109	Hs.108920	HT018 protein	2.18	
	110824	AI767183	Hs.26942	ESTs	12.20	
80	110776	AB032417	Hs.19545	fritzzled (Drosophila) homolog 4	1.75	
	110576	H60869	Hs.37889	ESTs	13.00	
	110369	AK000768	Hs.107872	hypothetical protein FLJ20761	5.60	
	110099	R44557	Hs.23748	ESTs	2.31	
	109984	AI796320	Hs.10299	Homo sapiens cDNA FLJ13545 fis, clone PL		
85	109958	AA001266	Hs.133521	ESTs	11.25	
	109893	AA884208	Hs.30484	ESTs	2.68	

5	109842	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	23.83		
	109837	H00656	Hs.29792	ESTs, Weakly similar to 138022 hypothe		17.20	3.91
	109796	AJ800515	Hs.12024	ESTs		9.60	
	109688	R41900	Hs.22245	ESTs			
	109648	H17800	Hs.7154	ESTs	22.80		
10	109613	H47315	Hs.27519	ESTs			
	109550	AW021488	Hs.26981	ESTs			
	109523	AW193342	Hs.24144	ESTs			1.89
	109472	AK001989	Hs.91165	hypothetical protein		6.00	
	109355	AA524525	Hs.48297	DKFZP586C1620 protein	15.00		
15	109260	AW978515	Hs.131915	KIAA0853 protein	25.60		
	108781	AA128654		gb:zn98g07.s1 Stratagene fetal retina 93	14.20		
	108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypothe	11.00		
	108573	AA086005		gb:z184c04.s1 Stratagene colon (937204)	26.00		
	108480	AL133092	Hs.68055	hypothetical protein DKFZp434i0428			
20	108382	NM_006770	Hs.67726	macrophage receptor with collagenous str			1.83
	108174	AA056632	Hs.303070	ESTs	15.20		
	108138	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp554G112 (fr		3.60	
	108087	AA045708	Hs.40545	ESTs	15.44		
	108048	AI797341	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT		11.40	
25	108041	AW204712	Hs.61957	ESTs			
	107997	AL049176	Hs.82223	chordin-like		4.76	
	107994	AA036811	Hs.48469	LIM domains containing 1			
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	14.20		
	107681	BE379594	Hs.49136	ESTs, Moderately similar to ALU7_HUMAN A	51.80		
30	107666	AA010611	Hs.60418	EST	29.20		
	107332	T87750	Hs.183297	DKFZP566F2124 protein	10.73		
	107292	BE166479	Hs.4789	Homo sapiens serologically defined breas	32.00		
	107230	AI034467	Hs.34650	ESTs	17.40		
	107168	W57578	Hs.237955	RAB7, member RAS oncogene family	10.43		
35	107160	AA314490	Hs.27669	KIAA1563 protein	11.40		
	107054	AI076459	Hs.15978	KIAA1272 protein			
	107029	AF264750	Hs.288971	myeloid/lymphoid or mixed-lineage leukem	21.40		
	106999	H93281	Hs.10710	hypothetical protein FLJ20417	35.80		
	106954	AF128847	Hs.204038	indolethylamine N-methyltransferase			1.76
40	106870	AI983730	Hs.26530	serum deprivation response (phosphatidyl			
	106865	AW192535	Hs.19479	ESTs	13.40		
	106844	AA485055	Hs.158213	sperm associated antigen 6		7.13	
	106820	NM_016831	Hs.12592	period (Drosophila) homolog 3		7.00	
	106818	AK002135	Hs.3542	hypothetical protein FLJ11273	13.00		
45	106797	AI768801	Hs.169943	Homo sapiens cDNA FLJ13569 fis, clone PL			2.05
	106773	AA478109	Hs.188833	ESTs			
	106747	NM_007118	Hs.171957	triple functional domain (PTPRF interact	12.60		
	106743	BE613328	Hs.21938	hypothetical protein FLJ12492	10.60		
	106667	AW360847	Hs.16578	ESTs			
50	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr			2.40
	106567	AW450408	Hs.86412	chromosome 9 open reading frame 5			1.78
	106562	AL031846	Hs.152151	plakophilin 4			1.76
	106536	AA329648	Hs.23804	ESTs, Weakly similar to PN0099 son3 prot			2.19
	106533	AL134708	Hs.145998	ESTs	23.20		
55	106507	AA259068	Hs.267819	protein phosphatase 1, regulatory (inhib	15.20		
	106490	AA404265	Hs.115537	putative dipeptidase			
	106474	BE383668	Hs.42484	hypothetical protein FLJ10618	10.44		
	106211	AA428240	Hs.126083	ESTs		29.80	
	105986	AB037722	Hs.8707	KIAA1301 protein		3.70	
60	105894	AI904740	Hs.25691	receptor (calcitonin) activity modifying			1.94
	105847	AW964490	Hs.32241	ESTs, Weakly similar to S65557 alpha-1C-			1.75
	105803	AW747996	Hs.160999	ESTs, Moderately similar to A56194 throm			2.47
	105731	AA834664	Hs.29131	nuclear receptor coactivator 2	10.71		
	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds			
65	105688	AI299139	Hs.17517	ESTs	23.40		
	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	37.20		
	105101	H63202	Hs.38163	ESTs		8.30	
	104989	R65998	Hs.285243	hypothetical protein FLJ22029		8.09	
	104986	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1			1.92
70	104969	AI670947	Hs.78406	phosphatidylinositol-4-phosphate 5-kinas		5.40	
	104903	AA36323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,		7.60	
	104896	AWD15318	Hs.23165	ESTs	13.80		
	104865	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C			
	104825	AA035613	Hs.141883	ESTs			1.87
75	104781	AA099904	Hs.21610	DKFZP434B203 protein			1.93
	104776	AA026349		gb:zj99i01.s1 Soares_pregnant_uterus_NbH	10.20		
	104691	U29590	Hs.37744	Homo sapiens beta-1 adrenergic receptor		5.69	
	104667	AI239923	Hs.30098	ESTs		3.82	
	104404	H58762		gb:EST00057 HE6W Homo sapiens cDNA clone	4.20		
80	104392	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	27.20		
	104212	AB002298	Hs.173035	KIAA0300 protein			1.91
	104074	AL162039	Hs.31422	Homo sapiens mRNA; cDNA DKFZp434M229 (fr	11.20		
	103749	AL135301	Hs.8768	hypothetical protein FLJ10849	10.86		
	103645	AW246253	Hs.7043	succinate-CoA ligase, GDP-forming, alpha	12.00		
85	103554	AI878826	Hs.323469	caveolin 1, caveolae protein, 22kD			1.80
	103541	AI815601	Hs.79197	CD83 antigen (activated B lymphocytes, I			
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2			
	103428	BE383507	Hs.78921	A kinase (PKA) anchor protein 1	11.20		
	103353	X89399	Hs.119274	RAS p21 protein activator (GTPase activa	19.80		

5	103295	X81479	Hs.2375	egf-like module containing, mucin-like,	3.60	
	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula		1.76
	103100	NM_005574	Hs.184585	LIM domain only 2 (rhombotin-like 1)		2.15
	103025	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t		
	102698	M18657	Hs.1867	progastricsin (pepsinogen C)		
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein	11.00	
	102580	U60808	Hs.152981	CDP-diacylglycerol synthase (phosphatida	25.40	
	102417	AA034127	Hs.153487	signal transducing adaptor molecule (SH3	14.00	
10	102363	NM_003734	Hs.198241	amine oxidase, copper containing 3 (vasc		
	102302	AA306342	Hs.69171	protein kinase C-like 2	10.86	
	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11		
	102188	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	7.40	
	102151	T27013	Hs.3132	steroidogenic acute regulatory protein	16.40	
15	101957	L28824	Hs.74101	spleen tyrosine kinase	15.40	
	101842	M93221	Hs.75182	mannose receptor, C type 1		
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant		
	101764	AI198550	Hs.81256	S100 calcium-binding protein A4 (calcium		1.78
	101716	AF050658	Hs.2563	tachykinin, precursor 1 (substance K, su	18.80	
20	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a)		2.22
	101447	M21305		gb:Human alpha satellite and satellite 3	504.80	
	101383	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	31.00	
	101346	AI738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N		1.75
	101345	NM_005795	Hs.152175	calcitonin receptor-like		
25	101336	NM_005732	Hs.75678	FBJ murine osteosarcoma viral oncogene h		2.24
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do		
	101277	BE297626	Hs.296049	microfibrillar-associated protein 4		
	101262	L35854		gb:Human dystrophin (dp140) mRNA, 5' end	19.00	
	101168	NM_005308	Hs.211559	G protein-coupled receptor kinase 5		2.01
30	101102	NM_003243	Hs.79059	transforming growth factor, beta recepto		
	101088	X70697	Hs.553	solute carrier family 6 (neurotransmitte	7.52	
	101068	AW970254	Hs.889	Charot-Leyden crystal protein	19.38	
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		1.91
	100893	BE245294	Hs.180789	S164 protein	15.40	
35	100770	W25797.comp	Hs.177486	amyloid beta (A4) precursor protein (pro	11.20	
	100716	X89887	Hs.172350	HIR (histone cell cycle regulation defec	14.80	
	100555	M69181		gb:Human nonmuscle myosin heavy chain-B	33.00	
	100425	NM_014747	Hs.78748	KIAA0237 gene product	16.20	
40	100408	D86640	Hs.56045	src homology three (SH3) and cysteine ri	4.00	
	100382	D83407	Hs.156007	Down syndrome critical region gene 1-lik	4.24	
	100351	D64158			6.20	
	100299	D49493	Hs.2171	growth differentiation factor 10	21.20	
	100134	AA305746	Hs.49	macrophage scavenger receptor 1		
	100108	U09577	Hs.76873	hyaluronoglucosaminidase 2		1.79
45	100095	Z97171	Hs.78454	myoclin, trabecular meshwork inducible	5.40	
	100066				11.29	

TABLE 39 shows the accession numbers for those primekeys lacking unigenelD's for Table 3A. For each probe set we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probe set identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
123619	371681_1	AA602964 AA609200
126433	127143_1	AA325506 AA099517 N89423
125831	1522905_1	H04043 D60988 D60337
126816	122973_1	AA248234 AA090985
126852	136135_1	AA399961 AA128347
121059	273450_1	AA393283 AA398628
120537	200885_1	AA811804 AA809404 AA286907 AW977624
122011	7617_2	AA431082
120934	177521_1	AA226198 AA226513 AA383773
123802	genbank_AA620448	AA620448
116814	genbank_H50834	H50834
118329	genbank_N63520	N63520
104404	H58762_at	H58762
104776	genbank_AA026349	AA026349
113502	genbank_T89130T89130	
101262	entrez_L35854	L35854
108573	genbank_AA086005	AA086005
101447	entrez_M21305	M21305
124357	genbank_N22401	N22401
108781	genbank_AA128654	AA128654
112794	genbank_R97018	R97018
100351	entrez_D64158	D64158
100555	lgr_HT2245	M69181 M81105 U51039

Table 4A shows 202 genes up-regulated in samples from patients treated with chemotherapy or radiotherapy. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5 Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 R1: average of AI for samples from patients treated with chemotherapy or radiotherapy divided by the average of AI for normal lung samples.

	Pkey	ExAccn	UnigenelD	Unigene Title	R1
	100113	NM_001269	Hs.84746	chromosome condensation 1	27.20
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	20.60
15	100210	D26361	Hs.3104	KIAA0042 gene product	20.40
	100225	D28539	Hs.167185	glutamate receptor, metabotropic 5	20.60
	100269	NM_001949	Hs.1189	E2F transcription factor 3	29.40
	100438	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	23.50
	100877	X80821	Hs.27973	KIAA0874 protein	35.56
20	100893	BE245294	Hs.180789	S164 protein	43.40
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	21.80
	101447	M21305		gb:Human alpha satellite and satellite 3	193.60
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	38.40
	101724	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	198.80
25	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	78.60
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	162.20
	101879	AA176374	Hs.243886	nuclear autoantigenic sperm protein (his	50.00
	101915	AF207881	Hs.155185	cytosolic ovarian carcinoma antigen 1	26.00
	101973	U41514	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:poly	37.20
30	102025	U04045	Hs.78934	mutS (E. coli) homolog 2 (colon cancer,	
	102031	U04898	Hs.2156	RAR-related orphan receptor A	32.00
	102052	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	51.20
	102391	AA296874	Hs.77494	deoxyguanosine kinase	13.90
	102420	U44060	Hs.14427	Homo sapiens cDNA: FLJ21800 fis, clone H	28.80
35	102610	U65011	Hs.30743	preferentially expressed antigen in mela	110.60
	102829	NM_006183	Hs.80962	neurotensin	116.80
	103090	NM_001975	Hs.146580	enolase 2, (gamma, neuronal)	2.30
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	181.40
	103507	AJ000512	Hs.296323	serum/glucocorticoid regulated kinase	49.20
40	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	86.60
	104660	BE298665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	42.60
	104896	AW015318	Hs.23165	ESTs	29.40
	105038	AW503733	Hs.9414	KIAA1488 protein	21.50
45	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	32.80
	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	20.20
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specifi	28.40
	106073	AL157441	Hs.17834	downstream neighbor of SON	25.40
	106205	AW965058	Hs.111583	ESTs, Weakly similar to I38022 hypotheti	32.00
50	106516	AL137311	Hs.234074	Homo sapiens mRNA; cDNA DKFZp761G02121 (40.60
	106533	AL134708	Hs.145998	ESTs	59.80
	106575	AW970602	Hs.105421	ESTs	43.40
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	50.80
	106851	A458823		gb:tk04g09.x1 NCI_CGAP_Lu24 Homo sapiens	53.40
55	106995	AB023139	Hs.37892	KIAA0922 protein	20.88
	107332	T87750	Hs.183297	DKFZP566F2124 protein	23.60
	107532	AA443473	Hs.173684	Homo sapiens mRNA; cDNA DKFZp762G207 (fr	57.20
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	49.00
	108609	BE409857	Hs.69499	hypothetical protein	19.67
60	108780	AU076442	Hs.117938	collagen, type XVII, alpha 1	48.17
	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (raklines	59.20
	109260	AW978515	Hs.131915	KIAA0863 protein	28.60
	109280	AK001355	Hs.279610	hypothetical protein FLJ10493	22.80
	109292	AW975746	Hs.188662	KIAA1702 protein	
65	109384	AA219172	Hs.86849	ESTs	21.00
	109415	U80736	Hs.110826	trinucleotide repeat containing 9	31.60
	109445	AA232103	Hs.189915	ESTs	24.20
	109502	AW967069	Hs.211556	hypothetical protein MGC5487	21.40
	109633	AW003785	Hs.170267	ESTs	20.40
70	109786	AI989482	Hs.146286	kinesin family member 13A	19.60
	109958	AA001266	Hs.133521	ESTs	24.00
	110920	N47224	Hs.20521	HMT1 (hnRNP methyltransferase, S. cerevi	28.40
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	36.00
	111084	H44186	Hs.15456	PDZ domain containing 1	61.20
	111132	AB037807	Hs.83293	hypothetical protein	24.60
75	111229	AW389845	Hs.110855	ESTs	27.20
	111337	AA837396	Hs.263925	LIS1-interacting protein NUDE1, rat homo	48.00
	111987	NM_015310	Hs.6763	KIAA0942 protein	37.80
	112046	AA383343	Hs.22116	CDC14 (cell division cycle 14, S. cerevi	26.80
	112268	W39509	Hs.22003	solute carrier family 6 (neurotransmitte	63.80
80	112685	R87650	Hs.33439	ESTs, Weakly similar to ALU1_HUMAN ALU	26.40
	112871	AL110216	Hs.12285	ESTs, Weakly similar to I55214 salivary	47.64
	112897	AW206463	Hs.3782	ESTs	22.00
	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	65.00
85	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	42.00
	113073	N39342	Hs.103042	microtubule-associated protein 1B	55.40

	113494	T91451	Hs.86538	ESTs	22.80
	113560	T91015	Hs.268626	ESTs	22.80
	113849	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	51.80
5	113950	A1267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (f	28.20
	114339	AA782845	Hs.22790	ESTs	20.20
	114365	H42169	Hs.18653	hypothetical protein FLJ14627	21.00
	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	25.80
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	23.60
10	114824	AA560961	Hs.305953	zinc finger protein 83 (HPF1)	27.20
	114837	BE244930	Hs.166895	ESTs	30.20
	114974	AW966931	Hs.179662	nucleosome assembly protein 1-Eke 1	20.80
	115075	AA814043	Hs.88045	ESTs	30.60
	115084	BE383668	Hs.42484	hypothetical protein FLJ10618	28.86
15	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	38.00
	115313	AA808001	Hs.184411	albumin	22.60
	115697	D31382	Hs.63325	transmembrane protease, serine 4	173.60
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	27.77
	116090	A1591147	Hs.61232	ESTs	20.80
20	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	164.20
	116399	AA889120	Hs.110637	homeo box A10	38.00
	117099	H93699	gb:yy16a11.s1 Soares fetal liver spleen	21.60	
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	49.40
	118091	AW005054	Hs.47883	ESTs, Weakly similar to KCC1_HUMAN CALCI	22.40
25	118138	AA374756	Hs.93560	Homo sapiens mRNA for KIAA1771 protein,	22.00
	118720	N73515	gb:za49d07.s1 Soares fetal liver spleen	20.00	
	118873	A1824009	Hs.44577	ESTs	19.40
	119126	R45175	Hs.117183	ESTs	111.20
	119717	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	33.00
30	119940	AL050097	Hs.272531	DKFZP586B0319 protein	31.00
	120266	A1807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	20.20
	120515	AA258355		gb:zr59c10.s1 Soares_NhHMPu_S1 Homo sapi	25.00
	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	95.40
	120983	AA398209	Hs.97587	EST	105.20
35	121054	AW976570	Hs.97387	ESTs	38.80
	121369	AW450737	Hs.128791	CGI-09 protein	41.60
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	30.80
	122612	AA974832	Hs.128708	ESTs	19.60
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H	33.20
40	123440	AI733692	Hs.112488	ESTs	23.17
	123596	AA421130	Hs.112640	EST	23.00
	123619	AA602864		gb:no97c02.s1 NCI_CGAP_Pr2 Homo sapiens	28.80
	124006	AI147155	Hs.270016	ESTs	77.60
	124169	BE079334	Hs.271630	ESTs	22.20
45	124281	AI333756	Hs.111801	arsenate resistance protein ARS2	42.20
	124472	N52517	Hs.102670	EST	32.60
	124617	AW628168	Hs.152684	ESTs	21.80
	124631	NM_014053	Hs.270594	FLVCR protein	30.40
	124839	R55784	Hs.140942	ESTs	21.20
50	125186	AA610620	Hs.181244	major histocompatibility complex, class	42.80
	125321	T86652	Hs.178294	ESTs	27.00
	125535	NM_013243	Hs.22215	secretogranin III	23.80
	125646	AA628962	Hs.75209	protein kinase (cAMP-dependant, catalyti	23.20
	125684	AW589427	Hs.158849	Homo sapiens cDNA: FLJ21663 fis, clone C	21.20
55	125724	AL360190	Hs.295978	Homo sapiens mRNA full length insert cDN	48.80
	125847	AW161885	Hs.249034	ESTs	31.00
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	21.20
	126077	M78772	Hs.210836	ESTs	49.80
	126299	AW979155	Hs.298275	amino acid transporter 2	21.80
60	126395	AA68004	Hs.278956	hypothetical protein FLJ12929	71.00
	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	23.20
	126509	R47400	Hs.23850	ESTs	23.80
	126538	AB030656	Hs.17377	coronin, actin-binding protein, 1C	23.10
	126666	AA648886	Hs.151999	ESTs	36.00
65	126812	AB037860	Hs.173933	nuclear factor I/A	20.80
	126872	AW450979		gb:U1-H-B13-ata-a-12-0-U1.s1 NCI_CGAP_Su	46.29
	127046	AA321948	Hs.293968	ESTs	22.80
	127431	AW771958	Hs.175437	ESTs, Moderately similar to PC4259 ferri	30.00
	127489	AA650250	Hs.272076	ESTs	20.80
70	127521	AW297206	Hs.164018	ESTs	25.20
	127742	AW293496	Hs.180138	ESTs	28.00
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	21.20
	127930	AA809572	Hs.123304	ESTs	20.54
	127968	AA830201	Hs.124347	ESTs	28.20
75	127987	AI022103	Hs.124511	ESTs	19.60
	128116	H07103	Hs.286014	Homo sapiens, clone IMAGE:3867243, mRNA	20.40
	128609	NM_003616	Hs.102455	survival of motor neuron protein interac	34.40
	128777	AI878918	Hs.10526	cysteine and glycine-rich protein 2	53.80
	128949	AA009647	Hs.8850	a disintegrin and metalloproteinase doma	23.00
80	129168	AI132888	Hs.109052	chromosome 14 open reading frame 2	37.60
	129404	AI267700	Hs.317584	ESTs	28.60
	129527	AA769221	Hs.270847	delta-tubulin	40.80
	129574	AA026815	Hs.11463	UMP-CMP kinase	31.20
	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	29.60
85	129785	H19006	Hs.184780	ESTs	72.20
	129970	AV655806	Hs.296198	chromosome 12 open reading frame 4	22.20

	130149	AW067805	Hs.172665	methylene tetrahydrofolate dehydrogenase	29.60
	130199	Z48579	Hs.172028	a disintegrin and metalloproteinase domain	27.60
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	28.36
5	130466	W19744	Hs.180059	Homo sapiens cDNA FLJ20653 fis, clone KA	20.20
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	22.40
	130617	M90516	Hs.1874	glutamine-fructose-6-phosphate transamin	19.60
	130703	R77776	Hs.18103	ESTs	19.40
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	21.40
10	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	110.00
	131028	AI879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	25.20
	131086	AL035461	Hs.2281	chromogranin B (secretogranin 1)	40.60
	131284	NM_001429	Hs.25272	E1A binding protein p300	24.60
	131775	AB014548	Hs.31921	KIAA0648 protein	21.00
15	131860	BE383676	Hs.334	Rho guanine nucleotide exchange factor (33.40
	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	60.80
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	20.40
	132084	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)	29.40
	132389	AA310393	Hs.190044	ESTs	32.40
20	132437	AA152106	Hs.4859	cyclin L, alpha-6a	27.40
	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	75.60
	132617	AF037335	Hs.6338	carbonic anhydrase XII	31.36
	132632	AJ076916	Hs.5398	guanine monophosphate synthetase	32.40
	132672	W27721	Hs.54697	Cdc42 guanine exchange factor (GEF) 9	23.40
25	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypotheti	61.20
	132771	Y10275	Hs.56407	phosphoserine phosphatase	22.33
	133070	U92649	Hs.64311	a disintegrin and metalloproteinase domain	23.50
	133153	AF070592	Hs.66170	HSPM-B protein	30.00
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	23.80
30	133282	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	51.60
	133350	AI499220	Hs.71573	hypothetical protein FLJ10074	33.00
	133592	AV652066	Hs.75113	general transcription factor IIIA	82.00
	133658	AA319146	Hs.75426	secretogranin II (chromogranin C)	
	133865	AB011155	Hs.170290	discs, large (Drosophila) homolog 5	69.33
35	134032	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	33.20
	134125	NM_014781	Hs.50421	KIAA0203 gene product	31.60
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	30.60
	134321	BE538082	Hs.8172	ESTs, Moderately similar to A45010 X-fin	23.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	49.20
40	134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	20.20
	134753	NM_006482	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl	20.80
	135002	AA448542	Hs.251677	G antigen 7B	37.60
	135029	H58818	Hs.187579	hydroxysteroid (17-beta) dehydrogenase	53.40
	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	31.60
45	135345	X53655	Hs.99171	neurotrophin 3	28.80

TABLE 4B shows the accession numbers for those primekeys lacking unique IDs for Table 4A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

PKkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

PKkey	CAT number	Accessions
123619	371681_1	AA602964 AA609200
126433	127143_1	AA325606 AA099517 N89423
126872	142696_1	AW450979 AA136653 AA136656 AW419381 AA984358 AA492073 BE168945 AA809054 AW238038 BE011212 BE011359 BE011367 BE011368 BE011362 BE011215 BE011365 BE011363
106851	322947_1	AA58623 AA639708 AA485409 R22065 AA485570
118720	genbank_N73515	N73515
120515	genbank_AA258356	AA258356
117099	321871_1	H93699 H97976 H80036
101447	entrez_M21305	M21305
123130	genbank_AA487200	AA487200

Table 5A shows 680 genes up-regulated in squamous cell carcinoma or adenocarcinoma lung tumors relative to normal lung and chronically diseased lung. These genes were selected from 59680 probesets on the Eos/Alfymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number							
	ExAccn:	Exemplar Accession number, Genbank accession number							
	UnigeneID:	Unigene number							
	Unigene Title:	Unigene gene title							
10	R1:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R2:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R3:	80th percentile of AI squamous cell carcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R4:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 80th percentile of AI for squamous cell carcinoma lung tumor samples.							
15	R5:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by 90th percentile of AI for normal and chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples							
20	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5
	100035			AFFX control: GAPDH					6.76
	100036			AFFX control: GAPDH					5.77
	100037			AFFX control: GAPDH					5.75
25	100071	A28102		Human GABAA receptor alpha-3 subunit		8.00			
	100114	X02308	Hs.82962	thymidylate synthetase					5.71
	100154	H60720	Hs.81892	KIAA0101 gene product	3.84				
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	3.33				
	100188	AW247090	Hs.57101	minichromosome maintenance deficient (S.					4.52
	100202	BE294407	Hs.99910	phosphofructokinase, platelet					5.49
30	100216	AA489908	Hs.1390	proteasome (prosome, macropain) subunit,					5.67
	100269	NM_001949	Hs.1189	E2F transcription factor 3	2.55				
	100287	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e					5.66
	100297	AU077258	Hs.182429	protein disulfide isomerase-related prot					3.81
	100330	AW410976	Hs.77152	minichromosome maintenance deficient (S.					4.50
35	100335	AW247529	Hs.6793	platelet-activating factor acetylhydrola	5.07				
	100360	W70171	Hs.75939	uridine monophosphate kinase					4.82
	100372	NM_014791	Hs.184339	KIAA0175 gene product					3.79
	100474	NM_000699	Hs.300280	amylase, alpha 2A; pancreatic				15.65	
40	100486	T19006	Hs.10842	RAN, member RAS oncogene family					5.49
	100491	D56165	Hs.275163	non-metastatic cells 2, protein (NM23B)					4.17
	100516	D90278	Hs.11	carcinoembryonic antigen-related cell ad		7.20			
	100522	X51501	Hs.99949	prolactin-induced protein				14.20	
	100559	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys	3.10				
	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid				9.30	
45	100629	AA015693	Hs.21291	mitogen-activated protein kinase kinase				20.60	
	100661	BE623001	Hs.132748	Homo sapiens ribosomal protein L39 mRNA,	3.85				
	100677	AA353686	Hs.57813	zinc ribbon domain containing, 1		8.60			
	100696	D14887	Hs.121686	general transcription factor IIA, 1 (37k				10.00	
	100709	N26539	Hs.100469	myeloid/lymphoid or mixed-lineage leukem			24.80		
50	100761	BE208491	Hs.295112	KIAA0618 gene product		7.60			
	100830	AC004770	Hs.4756	flap structure-specific endonuclease 1					7.99
	100867	U14622		gb:Human transketolase-like protein gene		10.20			
	100902	M16029	Hs.287270	ret proto-oncogene (multiple endocrine n		8.00			
	100906	AU076916	Hs.5398	guanine monophosphate synthetase					5.16
55	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	2.57				
	101045	J05614		gb:Human proliferating cell nuclear anti					4.69
	101051	NM_000175	Hs.180532	glucose phosphate isomerase					4.19
	101071	L02840	Hs.84244	potassium voltage-gated channel, Shab-re		12.91			
	101124	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	3.12				
60	101175	U82671	Hs.36980	melanoma antigen, family A, 2	3.50				
	101181	BE262621	Hs.73798	macrophage migration inhibitory factor (5.69
	101204	L24203	Hs.82237	ataxia-telangiectasia group D-associated	4.08				
	101210	L29301	Hs.2353	opioid receptor, mu 1			6.40		
	101216	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	2.53				
65	101228	AA333387	Hs.82916	chaperonin containing TCP1, subunit 6A (7.90
	101233	AL135173	Hs.878	sorbitol dehydrogenase					4.45
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	8.50				
	101342	U52112	Hs.182018	interleukin-1 receptor-associated kinase					4.17
	101346	AJ738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				21.89	
70	101389	NM_000892	Hs.1901	kallikrein B, plasma (Fletcher factor) 1				12.80	
	101396	BE267931	Hs.78996	proliferating cell nuclear antigen	3.24				
	101431	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)					7.90
	101448	NM_000424	Hs.195850	keratin 5 (epidermolysis bullosa simplex	8.31				
	101462	AL035668	Hs.73853	bone morphogenetic protein 2				38.80	
75	101466	BE262660	Hs.170197	glutamic-oxaloacetic transaminase 2, mit					4.01
	101484	AA053486	Hs.20315	interferon-induced protein with tetratr				12.00	
	101502	M26958		gb:Human parathyroid hormone-related pro	10.50				
	101505	AA307680	Hs.75692	asparagine synthetase					4.46
	101526	NM_002197	Hs.154721	aconitase 1, soluble	4.02				
80	101535	X57152	Hs.99853	fibrinogen					4.65
	101577	M34353	Hs.1041	v-ros avian UR2 sarcoma virus oncogene h				9.09	
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	54.00				
	101663	NM_003528	Hs.2178	H2B histone family, member Q	5.59				
	101664	AA436989	Hs.121017	H2A histone family, member A	7.00				
85	101669	L24498	Hs.80409	growth arrest and DNA-damage-inducible,		7.60			

5	101695	M69136	Hs.135626	chymase 1, mast cell	4.79		
	101724	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	15.21		
	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen)	55.50		
	101759	M80244	Hs.184601	solute carrier family 7 (cationic amino)		18.57	4.10
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant			
10	101804	M56699	Hs.169840	TTK protein kinase	4.50		
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	140.00		
	101833	AU076442	Hs.117938	collagen, type XVII, alpha 1	2.56		
	101842	M93221	Hs.75182	mannose receptor, C type 1		12.80	
	101851	BE260964	Hs.82045	midkine (neurite growth-promoting factor			5.88
15	102002	NM_002484	Hs.81469	nucleofold binding protein 1 (E.coli Min	7.80		
	102039	AL134223	Hs.306098	aldo-keto reductase family 1, member C1			4.35
	102072	U09410	Hs.78743	zinc finger protein 131 (clone pHZ-10)		7.40	
	102083	T35901	Hs.75117	interleukin enhancer binding factor 2, 4			5.12
	102111	L36198	Hs.81884	sulfotransferase family, cytosolic, 2A,		12.00	
20	102123	NM_001809	Hs.1594	centromere protein A (17kD)	6.20		
	102154	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kalinin	2.62		
	102193	AL036335	Hs.313	secreted phosphoprotein 1 (osteopontin,	5.85		
	102217	AA829978	Hs.301613	JTV1 gene			6.18
	102224	NM_002810	Hs.148495	proteasome (prosome, macropain) 26S subu			4.49
25	102234	AW163390	Hs.278554	heterochromatin-like protein 1			5.80
	102251	NM_004398	Hs.41706	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	4.50		
	102305	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)			5.15
	102330	BE288063	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta			4.17
	102340	U37055	Hs.278657	macrophage stimulating 1 (hepatocyte gro		9.33	
30	102348	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	8.87		
	102368	U39817	Hs.36820	Bloom syndrome	15.91		
	102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma		19.20	
	102404	NM_005429	Hs.79141	vascular endothelial growth factor C			14.00
	102537	U57094	Hs.50477	RAB27A, member RAS oncogene family			12.00
35	102581	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2			4.57
	102605	A135128	Hs.181369	ubiquitin fusion degradation 1-like			3.98
	102610	U65011	Hs.30743	preferentially expressed antigen in mela	77.50		
	102623	AW249285	Hs.37110	melanoma antigen, family A, 9	12.50		
	102642	AA205847	Hs.23016	G protein-coupled receptor		22.00	
40	102654	AV649989	Hs.24385	Human hbc647 mRNA sequence		12.00	
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein			12.80
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	6.50		
	102672	U72066	Hs.29287	retinoblastoma-binding protein 8	8.50		
	102687	NM_007019	Hs.93002	ubiquitin carrier protein E2-C			9.24
45	102696	BE540274	Hs.239	forkhead box M1			5.54
	102768	U82321		gb:Homo sapiens clone 14.9B mRNA sequenc	6.60		
	102781	BE258778	Hs.108809	chaperonin containing TCP1, subunit 7 (e			3.78
	102784	U85658	Hs.61795	transcription factor AP-2 gamma (activat			4.26
	102824	U90916	Hs.82845	Homo sapiens cDNA: FLJ21930 fis, clone H		14.40	
50	102829	NM_006183	Hs.80962	neurotensin	8.00		
	102888	A1346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1			5.50
	102892	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin		6.70	
	102913	NM_002275	Hs.80342	keratin 15	4.64		
	102935	BE561850	Hs.80506	small nuclear ribonucleoprotein polypept	2.93		
55	102951	X15218	Hs.2969	v-aki avian sarcoma viral oncogene homol		11.40	
	102983	BE387202	Hs.118638	non-metastatic cells 1, protein (NM23A)			7.26
	103023	AW500470	Hs.117950	multifunctional polypeptide similar to S	3.01		
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	27.90		
	103038	AA926960	Hs.334883	CDC28 protein kinase 1			8.79
60	103060	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin			4.27
	103099	A1693251	Hs.8248	NADH dehydrogenase (ubiquinone) Fe-S pro	9.80		
	103119	X63629	Hs.2677	cadherin 3, type 1, P-cadherin (placenta	4.05		
	103168	X53463	Hs.2704	glutathione peroxidase 2 (gastrointestin	3.07		
	103185	NM_006825	Hs.74368	transmembrane protein (63kD), endoplasm			5.62
65	103192	M22440	Hs.170009	transforming growth factor, alpha	7.40		
	103223	BE275607	Hs.1708	chaperonin containing TCP1, subunit 3 (g			4.70
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o		100.00	
	103316	X83301	Hs.324728	SMA5			9.80
	103375	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	9.71		
70	103376	AL036166	Hs.323378	coated vesicle membrane protein	14.00		
	103385	NM_007069	Hs.37189	similar to rat HREV107			11.00
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	2.93		
	103404	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,			5.15
	103430	BE564090	Hs.20716	translocase of inner mitochondrial membr			3.98
75	103446	X98834	Hs.79971	sal (Drosophila)-like 2		21.40	
	103476	Y07701	Hs.293007	aminopeptidase puromycin sensitive	13.00		
	103477	AJ011812	Hs.119018	transcription factor NRF		6.40	
	103478	BE514982	Hs.38991	S100 calcium-binding protein A2	5.02		
	103515	Y10275	Hs.56407	phosphoserine phosphatase	10.50		
80	103558	BE616547	Hs.2785	keratin 17	6.41		
	103580	AA328046	Hs.46405	polymerase (RNA) II (DNA directed) polyp			3.84
	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	78.50		
	103594	A1368580	Hs.816	SRY (sex determining region Y)-box 2	6.51		
	103636	NM_006235	Hs.2407	POU domain, class 2, associating factor	3.50		
85	103768	AF086009		gb:Homo sapiens full length insert cDNA			4.48
	103841	AA314821	Hs.38178	hypothetical protein FLJ23458	8.00		
	103847	AF219945	Hs.102237	tubby super-family protein	10.40		
	103913	AW967500	Hs.133543	ESTs			15.60
	104094	AA418187	Hs.330515	ESTs		6.60	

WO 02/086443

PCT/US02/12476

5	104150	AL122044	Hs.331633	hypothetical protein DKFZp566N034	26.00	
	104257	BE560621	Hs.9222	estrogen receptor binding site associate	6.80	3.98
	104261	AW248364	Hs.5409	RNA polymerase I subunit		
	104331	AB040450	Hs.279862	cdk inhibitor p21 binding protein	6.80	
	104415	BE410992	Hs.258730	heme-regulated initiation factor 2-alpha	10.29	
	104558	R56678	Hs.88959	hypothetical protein MGC4816	4.21	
	104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m		15.79
	104658	AA360954	Hs.27268	Homo sapiens cDNA: FLJ21933 fis, clone H		17.40
10	104660	BE298665	Hs.14846	Homo sapiens mRNA: cDNA DKFZp564D016 (tr	6.40	
	104689	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr		6.55
	104754	AJ206234	Hs.155924	cAMP responsive element modulator		10.00
	104758	BE560269	Hs.7010	NPD002 protein		4.47
	104971	BE311926	Hs.15830	hypothetical protein FLJ12691	2.87	
	105011	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	3.83	
15	105012	AF098158	Hs.9329	chromosome 20 open reading frame 1	2.86	
	105026	AA809485	Hs.124219	hypothetical protein FLJ12934	11.00	
	105076	AJ598252	Hs.37810	hypothetical protein MGC14833		5.01
	105132	AA148164	Hs.247280	HBV associated factor		3.99
	105143	AJ368836	Hs.24808	ESTs, Weakly similar to 138022 hypotheti	11.00	
20	105158	AW976357	Hs.234545	hypothetical protein NUF2R	16.00	
	105175	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	4.32	
	105200	AA328102	Hs.24641	cytoskeleton associated protein 2	3.00	
	105264	AA227934		gbzr57e08.s1 Soares_NhHMPu_S1 Homo sapi		10.00
25	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	3.69	9.20
	105409	AW505076	Hs.301855	DiGeorge syndrome critical region gene 8		
	105460	AW296078	Hs.271721	Homo sapiens, clone IMAGE:4179986, mRNA,	7.80	
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specif	4.12	
	105743	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	3.82	
30	105782	H09748	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro		27.00
	105848	AW954064	Hs.24951	ESTs	7.60	
	105891	U55984	Hs.289088	heat shock 90kD protein 1, alpha		4.14
	106019	AF221993	Hs.48743	McKusick-Kaufman syndrome	16.80	
	106069	BE566623	Hs.29899	ESTs, Weakly similar to G02075 transcrip	23.40	
35	106073	AL157441	Hs.17834	downstream neighbor of SON	9.50	
	106126	AA576953	Hs.22972	hypothetical protein FLJ13352	6.00	
	106159	AK001301	Hs.3487	hypothetical protein FLJ10439		3.95
	106220	D61329	Hs.32196	mitochondrial ribosomal protein L36		6.04
	106260	AJ097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S	13.20	
40	106300	Y10043	Hs.19114	high-mobility group (nonhistone chromoso		5.02
	106307	AA436174	Hs.37751	ESTs, Weakly similar to putative p150 [6.60	
	106318	AA025610	Hs.9605	cleavage and polyadenylation specific fa		5.04
	106341	AF191020	Hs.5243	hypothetical protein, estradiol-induced		7.25
	106440	AA449553	Hs.151393	glutamate-cysteine ligase, catalytic sub	13.80	
45	106481	D61594	Hs.17279	tyrosylprotein sulfotransferase 1	4.75	
	106586	AA243837	Hs.57787	ESTs		10.84
	106605	AW772298	Hs.21103	Homo sapiens mRNA: cDNA DKFZp564B076 (tr		45.60
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	28.00	
	106785	Y15227	Hs.20149	deleted in lymphocytic leukemia, 1	3.00	
50	106813	C05766	Hs.181022	CGI-07 protein		11.40
	106895	AK001826	Hs.25245	hypothetical protein FLJ11269	6.00	
	106913	AJ219346	Hs.86178	M-phase phosphoprotein 9	6.56	
	106919	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S		4.27
	107054	AJ076459	Hs.15978	KIAA1272 protein		34.80
55	107059	BE514410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	4.71	
	107098	AJ823593	Hs.27688	ESTs		24.80
	107104	AJ076640	Hs.15243	nucleolar protein 1 (120kD)		7.05
	107129	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.60	
	107198	AV657225	Hs.9946	KIAA1040 protein	19.20	
60	107203	D20426	Hs.41639	programmed cell death 2	7.60	
	107217	AL080235	Hs.35861	DKFZP586E1621 protein	9.50	
	107284	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	2.71	
	107318	T74445	Hs.5957	Homo sapiens clone 24416 mRNA sequence		8.71
	107516	X57152	Hs.99853	fibrillarin		4.33
65	107529	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)		4.00
	107728	AA019551	Hs.294151	Homo sapiens, clone IMAGE:3603836, mRNA,	10.80	
	107851	AA022953	Hs.61172	EST	8.00	
	107901	L42612	Hs.335952	keratin 6B	3.40	
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	2.88	
70	107932	AW392555	Hs.18878	hypothetical protein FLJ21620	7.50	
	108015	AW298357	Hs.49927	protein kinase NYD-SP15		23.40
	108056	AA043675	Hs.62633	ESTs		12.80
	108075	AJ887370	Hs.139709	hypothetical protein FLJ12572		12.80
	108187	BE245374	Hs.27842	hypothetical protein FLJ11210	7.00	
75	108295	N31255	Hs.161623	ESTs	6.60	
	108305	AA071391		gbzm61e05.r1 Stratagene fibroblast (937		11.80
	108393	AA075211		gbzm85a08.r1 Stratagene ovarian cancer		11.80
	108480	AL133092	Hs.68055	hypothetical protein DKFZp434I0428		20.80
	108554	AA084948		gbzn13b09.s1 Stratagene hNT neuron (937	6.40	
80	108573	AA086005		gbz284c04.s1 Stratagene colon (937204)		25.40
	108584	AA088326	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	9.60	
	108597	AK000292	Hs.278732	hypothetical protein FLJ20285		14.60
	108695	AB029000	Hs.70823	KIAA1077 protein	3.00	
	108699	AA121514	Hs.70832	ESTs		10.00
85	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c	11.00	
	108780	AJ076442	Hs.117938	collagen, type XVII, alpha 1	11.21	

5	108810	AW295647	Hs.71331	hypothetical protein MGC5350	8.50			
	108816	AA130884	Hs.270501	ESTs, Moderately similar to ALU2_HUMAN		7.40		
	108857	AK001458	Hs.62180	anillin (Drosophila Scraps homolog), act	4.00			
	108860	AA133334	Hs.129911	ESTs	6.09			
	108937	AL050107	Hs.24341	transcriptional co-activator with PDZ-bl	3.00			
10	109010	NM_007240	Hs.44229	dual specificity phosphatase 12	2.69			
	109121	BE388387	Hs.49767	NADH dehydrogenase (ubiquinone) Fe-S pro			4.53	
	109166	AA219691	Hs.73625	RAB6 Interacting, kinase-like (rabkines	10.58			
	109227	AA766998	Hs.85874	Human DNA sequence from clone RP11-16L21		9.00		
	109415	U80736	Hs.110826	trinucleotide repeat containing 9		51.40		
15	109418	AB66946	Hs.161707	ESTs			11.00	
	109454	AA232255	Hs.295232	ESTs, Moderately similar to A46010 X-II		17.60		
	109502	AW967069	Hs.211556	hypothetical protein MGC5487		9.49		
	109543	AA564994	Hs.222851	ESTs	12.67			
	109648	H17800	Hs.7154	ESTs			10.40	
20	109680	AB037734	Hs.4993	KIAA1313 protein		33.20		
	109700	F09609		gb:HSC33H092 normalized infant brain cDN			16.00	
	109704	AI743880	Hs.12876	ESTs		11.00		
	109792	R49625		gb:yg61f03.s1 Soares infant brain 1NIB H	4.00		12.60	
	109981	BE546208	Hs.26090	hypothetical protein FLJ20272				
25	109998	AL042201	Hs.21273	transcription factor NYD-sp10	7.80			
	110039	H11938	Hs.21907	histone acetyltransferase	7.00			
	110156	AA581322	Hs.4213	hypothetical protein MGC16207			4.24	
	110500	AA907723	Hs.36982	ESTs	4.50			
	110551	AW450381	Hs.14529	ESTs		8.60		
30	110561	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.06			
	110854	BE612992	Hs.27931	hypothetical protein FLJ10607 similar to		6.80		
	110886	AW274992	Hs.72249	three-PDZ containing protein similar to		8.80		
	110916	BE178102	Hs.24349	ESTs		6.80		
	111003	N52980	Hs.83765	dihydrofolate reductase			16.80	
35	111337	AA837396	Hs.263925	LIS1-Interacting protein NUDE1, rat homo	2.54			
	111434	R01608	Hs.142736	ESTs			9.80	
	111439	AI476429	Hs.19238	ESTs			10.40	
	111540	U82670	Hs.9786	zinc finger protein 275		15.40		
	111597	R11499	Hs.189716	ESTs			9.20	
40	111895	T80581	Hs.12723	Homo sapiens clone 25153 mRNA sequence	6.80			
	111929	AF027208	Hs.112360	prominin (mouse)-like 1			14.67	
	112054	R43590		gb:yc85g02.s1 Soares infant brain 1NIB H	10.80			
	112210	R49645	Hs.7004	ESTs			10.20	
	112244	AB029000	Hs.70823	KIAA1077 protein	2.99			
45	112382	R59904		gb:yc07g12.s1 Soares infant brain 1NIB H	6.60			
	112392	R60763	Hs.193274	ESTs, Moderately similar to I57588 HSrel		7.10		
	112442	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regl	3.00			
	112539	R70318	Hs.339730	ESTs			37.20	
	112772	AI992283	Hs.35437	ESTs, Moderately similar to I38026 MLN 6			14.60	
50	112869	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin				4.83
	112935	R71449	Hs.268760	ESTs	2.73			
	112970	AA694010	Hs.6932	Homo sapiens clone 23609 mRNA sequence			12.00	
	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	11.50			
	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f		10.89		
55	113063	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	15.00			
	113073	N39342	Hs.103042	microtubule-associated protein 1B		15.31		
	113078	T40444	Hs.118354	CAT56 protein	7.00			
	113238	R45467	Hs.189813	ESTs			41.20	
	113591	T91881	Hs.200597	KIAA0563 gene product			9.40	
60	113702	T97307		gb:ye53h05.s1 Soares fetal liver spleen	25.00			
	113844	AJ369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE			13.91	
	113984	R96696	Hs.35598	ESTs	7.80			
	114073	R44953	Hs.22908	Homo sapiens mRNA; cDNA DKFZp434J1027 (f	7.20			
	114162	AF155661	Hs.22265	pyruvate dehydrogenase phosphatase	3.42			
65	114208	AL049466	Hs.7859	ESTs		6.74		
	114251	H15261	Hs.21948	ESTs			33.20	
	114285	R44338	Hs.22974	ESTs			13.20	
	114313	H18456	Hs.27946	ESTs			10.00	
	114339	AA782845	Hs.22790	ESTs	7.80			
70	114407	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f			4.14	
	114560	AI452459	Hs.165221	ESTs		9.80		
	114699	AA127386		gb:zn90d09.r1 Stratagene lung carcinoma	7.60			
	114767	AI859865	Hs.154443	minichromosome maintenance deficient (S	3.21			
	114793	AA158245		gb:zo76c03.s1 Stratagene pancreas (93720		6.00		
75	114833	AI417215	Hs.87159	hypothetical protein FLJ12577			11.40	
	115047	BE270930	Hs.82916	chaperonin containing TCP1, subunit 6A (4.31	
	115060	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3			4.03	
	115097	AA256213	Hs.72010	ESTs			35.40	
	115113	AA256460		gb:zr81a04.s1 Soares_NhHMPu_S1 Homo sapi			15.20	
80	115123	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m				4.19
	115134	AW968073	Hs.194331	ESTs, Highly similar to A55713 inositol			12.40	
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	25.00			
	115347	AA356792	Hs.334824	hypothetical protein FLJ14825		7.00		
	115414	AA662240	Hs.283099	AF15q14 protein	3.25			
85	115522	BE614387	Hs.333893	c-Myc target JPO1	3.68			
	115536	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	10.50			
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1			24.40	
	115545	AI207410	Hs.69280	Homo sapiens, clone IMAGE:3636299, mRNA,	4.17			
	115548	AW016811	Hs.234478	Homo sapiens cDNA: FLJ22648 fis, clone H		6.00		

	115652	BE093589	Hs.38178	hypothetical protein FLJ23468	3.81				
	115697	D31382	Hs.63325	transmembrane protease, serine 4	62.14				
	115793	AA424883	Hs.70333	hypothetical protein MGC10753				11.80	
5	115816	BE042915	Hs.287588	Homo sapiens cDNA FLJ13675 fis, clone PL				9.71	
	115892	AA291377	Hs.50831	ESTs		27.40			
	115906	A1767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	2.53				
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	11.82				
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970				34.29	
10	115978	AL035864	Hs.69517	cDNA for differentially expressed CO16 g				8.23	
	115985	AA447709	Hs.268115	ESTs, Weakly similar to T08599 probable	3.00				
	116090	AI591147	Hs.61232	ESTs	5.17				
	116096	AA682382	Hs.59982	ESTs		8.20			
	116127	AF126743	Hs.279884	DNAJ domain-containing		10.60			
15	116157	BE439838	Hs.44298	mitochondrial ribosomal protein S17				5.82	
	116190	AI949095	Hs.57776	ESTs, Weakly similar to T22341 hypothet				4.08	
	116278	NM_003686	Hs.47504	exonuclease 1	9.50				
	116335	AK001100	Hs.41690	desmocollin 3	3.67				
	116496	AW450694	Hs.21433	hypothetical protein DKFZp547J036		7.00			
20	116503	AI925316	Hs.212617	ESTs				12.60	
	116674	A1768015	Hs.92127	ESTs		32.00			
	116929	AA586922	Hs.80475	polymerase (RNA) II (DNA directed) polyp	7.60				
	116973	AI702054	Hs.166982	phosphatidylinositol glycan, class F	9.80				
	116993	AI417023	Hs.40478	ESTs				10.20	
25	117079	H92325		gbys85f05.s1 Soares retina N2b4HR Homo				15.20	
	117317	AI263517	Hs.43322	ESTs				13.40	
	117326	N23629	Hs.241420	Homo sapiens mRNA for KIAA1756 protein,				20.60	
	117398	W20128	Hs.296039	ESTs				10.60	
	117412	N32536	Hs.42645	ESTs				16.00	
30	117519	N32528	Hs.146286	kinesin family member 13A				9.11	
	117693	AW179019	Hs.112110	mitochondrial ribosomal protein L42				4.01	
	117721	N46100	Hs.93939	EST				19.80	
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	2.71				
	117903	AA768283	Hs.47111	ESTs				17.80	
35	117992	AI015709	Hs.172089	Homo sapiens mRNA; cDNA DKFZp58612022 (f				10.60	4.17
	118013	AI674126	Hs.94031	ESTs			8.82		
	118017	AI813444	Hs.42197	ESTs					
	118186	N22886	Hs.42380	ESTs		7.00			
	118325	AI868065	Hs.166184	intersectin 2				13.80	
40	118367	N64269	Hs.48946	EST			6.14		
	118368	N64339	Hs.48956	gap junction protein, beta 6 (connexin 3	3.14				
	118472	AL157545	Hs.42179	bromodomain and PHD finger containing, 3			12.40		
	118709	AA232970	Hs.293774	ESTs				12.20	
	119025	BE003760	Hs.55209	Homo sapiens mRNA; cDNA DKFZp434K0514 (f	4.50				
45	119027	AF086161	Hs.114511	hypothetical protein FLJ11808	3.22				
	119052	R10889		gbv138d02.s1 Soares fetal liver spleen		9.60			
	119164	AF221993	Hs.46743	McKusick-Kaufman syndrome			6.60		
	119186	AI979147	Hs.101265	hypothetical protein FLJ22593				10.80	
	119243	T12603		gb:CHR90123 Chromosome 9 exon II Homo sa				9.44	
50	119490	AA195276	Hs.263858	ESTs, Moderately similar to B34087 hypot				11.80	
	119499	AI918906	Hs.55080	ESTs			14.80		
	119599	W45552		gb:zc26d03.s1 Soares_senescent_fibroblas		12.60			
	119780	NM_016625	Hs.191381	hypothetical protein	17.00				
	119845	W79123	Hs.58561	G protein-coupled receptor 87	13.50				
55	119941	AA699485	Hs.58896	ESTs		8.00			
	119994	AA642402	Hs.59142	ESTs		7.73			
	120102	W67353	Hs.170218	KIAA0251 protein			39.60		
	120104	AK000123	Hs.180479	hypothetical protein FLJ20116	2.91				
	120294	AK000059	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par			8.20		
60	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	8.73				
	120599	AA804448	Hs.104463	ESTs		7.00			
	120699	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos				10.00	
	120715	AA292700		gb:zs59a06.s1 NCL_CGAP_GCB1 Homo sapiens		9.40			
	120821	Y19062	Hs.96870	staurin (Drosophila, RNA-binding protein				13.80	
65	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol		9.00			
	120880	AA360240	Hs.97019	EST		15.60			
	120983	AA398209	Hs.97587	EST			27.66		
	121034	AL389951	Hs.271623	nucleoporin 50kD			20.80		
	121121	AA399371	Hs.189095	similar to SALL1 (sal (Drosophila)-like		22.80			
70	121313	AA402713	Hs.97872	ESTs				10.00	
	121369	AW450737	Hs.128791	CGI-09 protein	25.71				
	121376	AA448103	Hs.187958	solute carrier family 6 (neurotransmitte				5.42	
	121476	AA412311	Hs.97903	ESTs		8.30			
	121509	AA868939	Hs.97888	ESTs		8.59			
75	121553	AA412488	Hs.48820	TATA box binding protein (TBP)-associat	18.50				
	121753	AK000552	Hs.323518	WD repeat domain 5	7.00				
	121838	AA425680	Hs.98441	ESTs				10.40	
	121857	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	6.00				
	121991	AA430058	Hs.98649	EST				12.20	
80	122089	AW016543	Hs.98682	hypothetical protein FKSG32		8.60			
	122105	AW241685	Hs.98699	ESTs		6.14			
	122163	AA435702	Hs.98829	EST				10.40	
	122318	AA429743		gb:zv60b05.r1 Soares_testis_NHT Homo sap				18.20	
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	13.50				
85	122338	AA443311	Hs.98998	ESTs	4.80				
	122414	AI313473	Hs.99067	ESTs, Weakly similar to S47073 finger pr		8.00			

	122512	AF053305	Hs.98658	budding uninhibited by benzimidazoles 1	8.80		
	122516	AA493352	Hs.99217	ESTs		9.40	
	122702	AJ220089	Hs.99439	ESTs	9.20		
	122852	AI580056	Hs.98992	ESTs		10.40	
5	122925	AW268962	Hs.111335	ESTs	6.80		
	123005	AW369771	Hs.52620	Integrin, beta 8		12.60	
	123044	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro			5.35
	123160	AA486687	Hs.284235	ESTs, Weakly similar to I38022 hypothe		6.06	
	123315	AA496369		gb:zv37d10.s1 Soares ovary tumor NbHOT H		12.40	
10	123329	Z47542	Hs.179312	small nuclear RNA activating complex, po		11.80	
	123497	AA765256	Hs.135191	ESTs, Weakly similar to unnamed protein	12.00		
	123518	AL035414	Hs.21068	hypothetical protein		13.00	
	123519	AW015887	Hs.112574	ESTs	12.20		
	123614	AK000492	Hs.98806	hypothetical protein		7.80	
15	123616	AA680003	Hs.109363	Homo sapiens cDNA: FLJ23603 fis, clone L			10.60
	123673	BE550112	Hs.158549	ESTs, Weakly similar to T2D3_HUMAN TRANS	23.00		
	123727	AI083986	Hs.282977	hypothetical protein FLJ13490	7.00		
	123731	AA609839		gb:ae62f01.s1 Stralagene lung carcinoma		9.80	
	123752	AA227714	Hs.179703	KIAA0129 gene product	3.50		
20	123800	AA621223	Hs.112953	EST		12.80	
	124006	AJ147155	Hs.270016	ESTs	97.00		
	124059	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypothe	3.02		
	124069	AF134160	Hs.7327	claudin 1		27.80	
25	124181	T96509	Hs.248549	ESTs, Moderately similar to S65657 alpha			35.80
	124273	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	7.20		
	124297	AL080215	Hs.102301	Homo sapiens mRNA: cDNA DKFZp586J0323 (f			11.00
	124305	AW963221		gb:EST375294 MAGE resequences, MAGH Homo			16.00
	124676	AJ360119.comp	Hs.181013	phosphoglycerate mutase 1 (brain)			6.08
30	124874	BE550182	Hs.127826	RaGEF-like protein 3, mouse homolog			21.00
	124904	AK000483	Hs.93872	KIAA1682 protein	9.40		
	124969	AI650360	Hs.100256	ESTs		10.80	
	125000	T58615	Hs.110640	ESTs		9.80	
	125201	AA693960	Hs.103158	ESTs, Weakly similar to T33296 hypothe	7.60		
35	125266	W90022	Hs.186809	ESTs, Highly similar to LCT2_HUMAN LEUKO	6.59		
	125299	T32982	Hs.102720	ESTs			9.57
	125356	AI057052	Hs.133554	ESTs, Weakly similar to Z195_HUMAN ZINC			14.00
	125370	AA256743	Hs.134158	Homo sapiens, Similar to KIAA0092 gene p	8.20		
	125418	AA777690	Hs.188501	ESTs		13.20	
40	125433	AL162066	Hs.54320	hypothetical protein DKFZp762D096	21.40		
	125437	AI609449	Hs.140197	ESTs	6.96		
	125446	BE219987	Hs.166982	phosphatidylinositol glycan, class F	8.80		
	125711	AA305800	Hs.5672	hypothetical protein AF140225		11.20	
	125756	BE174587	Hs.289721	growth arrest specific transcript 5			4.31
45	125757	AJ274906	Hs.166835	ESTs, Highly similar to 1814460A p53-ass			15.60
	125769	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	3.20		
	125839	AW836261	Hs.337717	ESTs	8.20		
	125850	W85858	Hs.99804	ESTs	2.65		
	125875	H14480		gb:ym18b09.r1 Soares infant brain 1NIB H	7.40		
50	125924	BE272506	Hs.82109	syndecan 1			4.23
	125972	AI927475	Hs.35406	ESTs, Highly similar to unnamed protein			3.98
	126034	H60340		gb:yr39b04.r1 Soares fetal liver spleen		10.60	
	126327	AA432266	Hs.44648	ESTs	11.60		
	126345	N49713		gb:yv23f05.s1 Soares fetal liver spleen	6.67		
55	126435	AW614529	Hs.285847	CG-19 protein		10.60	
	126487	AA283809	Hs.184601	solute carrier family 7 (cationic amino			4.38
	126521	AI475110	Hs.203933	ESTs	6.60		
	126522	W31912		gb:zc76d03.s1 Pancreatic Islet Homo sapi		14.80	
60	126543	AL035864	Hs.69517	cDNA for differentially expressed CO16 g			4.01
	126567	AA058394	Hs.57887	ESTs, Weakly similar to KIAA0758 protein	7.80		
	126605	AA676910		gb:zj65h07.s1 Soares fetal liver spleen		11.60	
	126627	AA497044	Hs.20887	hypothetical protein FLJ10392		14.60	
	126628	N49776	Hs.170994	hypothetical protein MGC10946	8.00		
	126737	AW976516	Hs.283707	Homo sapiens cDNA: FLJ21354 fis, clone C	2.92		
	126795	AW975076	Hs.172589	nuclear phosphoprotein similar to S. cer	7.50		
65	126802	AW805510	Hs.97056	hypothetical protein FLJ21634	11.60		
	126892	AF121856	Hs.284291	sorting nexin 6	3.50		
	126928	AA480902	Hs.137401	ESTs		22.83	
	126979	AA210954		gb:zc89h10.r1 Stralagene hNT neuron (937		11.80	
70	126986	AI279892	Hs.46801	sorting nexin 14		11.60	
	126992	AI809521		gb:w30e03.x1 Soares_NFL_T_GBC_S1 Homo s		20.80	
	127066	R25066		gb:yq42c07.r1 Soares infant brain 1NIB H		27.60	
	127099	AA347668		gb:EST54026 Fetal heart II Homo sapiens		21.60	
	127139	AA830233	Hs.293585	ESTs		11.20	
75	127209	AA305023	Hs.81964	SEC24 (S. cerevisiae) related gene famil	3.10		
	127221	BE062109	Hs.241551	chloride channel, calcium activated, fam	2.76		
	127225	AA315933	Hs.120879	ESTs		16.80	
	127313	AK002014	Hs.47546	Homo sapiens cDNA FLJ11458 fis, clone HE	14.00		
	127444	AW978474	Hs.7560	Homo sapiens mRNA for KIAA1729 protein,		13.60	
	127500	AW971353	Hs.162115	ESTs	11.20		
80	127524	AI243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin		7.80	
	127540	N45572	Hs.105362	Homo sapiens, clone MGC:18257, mRNA, com	3.53		
	127599	AA613204	Hs.150399	ESTs		13.80	
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture		28.00	
85	127662	W80755	Hs.8294	KIAA0196 gene product		18.80	
	127668	AI343257	Hs.139993	ESTs		11.20	

	127746	AU239495	Hs.120189	ESTs			14.18
	127812	AA741368	Hs.291434	ESTs	4.50		
	127817	AA836641	Hs.163085	ESTs			24.60
5	127959	AJ302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L			9.20
	127960	AI613226	Hs.41569	phosphatidic acid phosphatase type 2A			16.83
	127969	F06498	Hs.93748	Homo sapiens cDNA FLJ14676 fis, clone NT	13.60		
	128015	Z21169	Hs.334659	hypothetical protein MGC14139	7.00		
	128027	AJ433721	Hs.164153	ESTs			37.40
	128077	AJ310330	Hs.128720	ESTs			9.60
10	128166	NM_006147	Hs.11801	interferon regulatory factor 6			9.24
	128226	AI284940	Hs.289082	GM2 ganglioside activator protein	19.00		
	128305	AI954968	Hs.279009	matrix Gla protein			10.40
	128341	AA191420	Hs.185030	ESTs	9.00		
	128527	AA504583	Hs.101047	transcription factor 3 (E2A immunoglobul			4.30
15	128539	R46163	Hs.258618	ESTs	12.60		
	128568	H12912	Hs.274691	adenylate kinase 3			4.56
	128572	AA933022	Hs.256583	interleukin enhancer binding factor 3, 9			10.00
	128777	AJ878918	Hs.10526	cysteine and glycine-rich protein 2		16.80	
	128781	N71826	Hs.105465	small nuclear ribonucleoprotein polypept			4.48
20	128796	AJ000152	Hs.105924	defensin, beta 2	8.12		
	128920	AA622037	Hs.166468	programmed cell death 5			4.62
	128924	BE279383	Hs.26557	plakophilin 3			4.04
	128971	H05132	Hs.107510	ESTs	12.60		
	129008	AL079648	Hs.301088	ESTs	8.80		
25	129041	BE382756	Hs.169902	solute carrier family 2 (facilitated glu			6.05
	129075	BE250162	Hs.83765	dihydrofolate reductase	2.59		
	129105	AI769160	Hs.108681	Homo sapiens brain tumor associated prot		6.67	
	129189	AB023179	Hs.9059	KIAA0952 protein	8.00		
	129229	AF013758	Hs.109543	polyadenylate binding protein-interactin	4.00		
30	129241	AJ878857	Hs.109706	hematological and neurological expressed			4.06
	129300	W94197	Hs.110165	ribosomal protein L26 homolog	2.55		
	129404	AI267700	Hs.317584	ESTs	18.00		
	129457	X61959	Hs.207776	aspartylglucosaminidase	6.50		
	129466	L42583	Hs.334309	keratin 6A	12.94		
35	129494	AI148976	Hs.112062	ESTs			11.00
	129505	AF061812	Hs.115947	keratin 16 (focal non-epidermolytic palm			4.46
	129641	AJ911627	Hs.11805	ESTs			12.00
	129665	AW163331	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic retic			4.70
40	129703	BE388665	Hs.179999	Homo sapiens, clone IMAGE:3457003, mRNA			4.02
	129720	AA156214	Hs.12152	APMCF1 protein			5.71
	129748	M16707	Hs.123053	H4 histone, family 2	3.50		
	129890	AJ858872	Hs.282804	hypothetical protein FLJ22704			4.21
	129896	BE295568	Hs.13225	UDP-Gal:beta-GlcNAc beta 1,4-galactosylt	2.56		
	129945	BE514376	Hs.165998	PAI-1 mRNA-binding protein			4.03
45	130010	AA301116	Hs.142838	nucleolar phosphoprotein Nopp34		7.00	
	130026	T40480	Hs.332112	EST	6.40		
	130080	X14850	Hs.147097	H2A histone family, member X			4.65
	130149	AW057805	Hs.172665	methyleneletrahydrofolate dehydrogenase	2.74		
50	130285	AA063546	Hs.75981	ubiquitin specific protease 14 (tRNA-gua		7.40	
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic			3.91
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	4.87		
	130500	AB007913	Hs.158291	KIAA0444 protein			9.60
	130524	U89995	Hs.159234	forkhead box E1 (thyroid transcription f		13.40	
	130541	X05608	Hs.211584	neurofilament, light polypeptide (68kD)		8.20	
55	130553	AF062649	Hs.252587	pituitary tumor-transforming 1			6.06
	130567	AA383032	Hs.1608	replication protein A3 (14kD)		7.00	
	130577	M69241	Hs.162	insulin-like growth factor binding prote	3.04		
	130627	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	3.87		
60	130648	AI458165	Hs.17296	hypothetical protein MGC2376			16.20
	130697	L29472	Hs.1802	major histocompatibility complex, class			17.80
	130744	H59696	Hs.18747	POP7 (processing of precursor, S. cerevi			5.28
	130800	AI187292	Hs.19574	hypothetical protein MGC5469			4.43
	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	16.84		
65	130869	J03626	Hs.2057	uridine monophosphate synthetase (orotat			4.92
	130925	AF093419	Hs.169378	multiple PDZ domain protein			9.60
	130994	W17044	Hs.327337	ESTs		12.40	
	131028	AJ879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	10.21		
	131031	NM_001650	Hs.288650	aquaporin 4			9.80
70	131041	T15767	Hs.22452	Homo sapiens mRNA for KIAA1737 protein,			9.60
	131058	W28545	Hs.101514	hypothetical protein FLJ10342			17.00
	131090	AI143139	Hs.2288	visinin-like 1	2.74		
	131112	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f		8.80	
	131148	AW953575	Hs.303125	p53-induced protein PIGPC1	3.12		
75	131185	BE280074	Hs.23960	cyclin B1	3.07		
	131200	BE540516	Hs.293732	hypothetical protein MGC3195	3.07		
	131219	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	2.87		
	131257	AW339037	Hs.24908	ESTs			14.67
	131375	AW293165	Hs.143134	ESTs		19.20	
80	131460	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.50		
	131476	AI521663	Hs.334644	hypothetical protein FLJ14668	15.00		
	131510	BE245374	Hs.27842	hypothetical protein FLJ11210		7.80	
	131646	BE302464	Hs.30057	MRS2 (S. cerevisiae)-like, magnesium hom		7.00	
	131786	BE000971	Hs.306083	Novel human gene mapping to chromosome 22	2.65		
85	131839	AB014533	Hs.33010	KIAA0633 protein			35.20
	131843	AA192315	Hs.184052	putative Rab5-interacting protein			4.11

	131877	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	19.00		
	131885	BE502341	Hs.3402	ESTs	6.48		
	131921	AA456093	Hs.34720	ESTs		8.40	
5	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	56.00		
	131958	NM_014062	Hs.3566	ART-4 protein			3.82
	131965	W79283	Hs.35962	ESTs	3.03		
	132000	AW247017	Hs.36978	melanoma antigen, family A, 3		9.80	
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	3.30		
10	132109	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagonist	21.00		
	132114	NM_006152	Hs.40202	lymphoid-restricted membrane protein		8.40	
	132162	AA315805	Hs.94560	desmoglein 2			12.25
	132164	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio	2.70		
	132180	NM_004460	Hs.418	fibroblast activation protein, alpha	2.71		
15	132181	AW961231	Hs.16773	Homo sapiens clone TCCCA00427 mRNA sequ	3.83		
	132182	NM_014210	Hs.70499	ecotropic viral integration site 2A			13.20
	132231	AA662910	Hs.42635	hypothetical protein DKFZp434K2435	9.50		
	132277	AK001745	Hs.184628	hypothetical protein FLJ10883	4.50		
	132328	NM_014787	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe			9.20
20	132394	AK001680	Hs.30488	DKFZP434F091 protein			19.80
	132424	AA417878	Hs.48401	ESTs, Moderately similar to ALU8_HUMAN A		8.60	
	132528	T78736	Hs.50758	SMC4 (structural maintenance of chromoso		27.40	
	132543	BE568452	Hs.5101	protein regulator of cytokinesis 1	4.38		
	132544	L19778	Hs.51011	H2A histone family, member P		7.00	
25	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	2.64		
	132552	BE621985	Hs.296922	thiopurine S-methyltransferase			15.83
	132581	AK000631	Hs.52256	hypothetical protein FLJ20624		6.60	
	132617	AF037335	Hs.5338	carbonic anhydrase XII	4.95		
	132638	AI796870	Hs.54277	DNA segment on chromosome X (unique) 992		8.20	
30	132653	Z15008	Hs.54451	laminin, gamma 2 (niclen (100kD), kalini	4.38		
	132669	W35866	Hs.293981	guanine nucleotide binding protein (G pr			4.36
	132710	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	4.60		
	132771	Y10275	Hs.56407	phosphoserine phosphatase	3.71		
	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,			9.48
35	132833	U78525	Hs.57783	eukaryotic translation initiation factor			5.83
	132892	AWB34050	Hs.9973	tensin		12.00	
	132906	BE613337	Hs.234896	germinin	3.09		
	132959	AW014195	Hs.61472	ESTs, Weakly similar to YAE6_YEAST HYPOT			3.87
	132962	AA576635	Hs.6153	CGI-48 protein	3.50		
40	132990	X77343	Hs.334334	transcription factor AP-2 alpha (activat	6.18		
	132994	AA112748	Hs.279905	clone HQ0310 PRO0310p1	3.19		
	133000	AL042444	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	2.95		
	133050	X73424	Hs.63788	propionyl Coenzyme A carboxylase, beta p	2.55		
	133083	BE244588	Hs.6456	chaperonin containing TCP1, subunit 2 (b			4.00
45	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso			8.96
	133134	AF198620	Hs.65648	RNA binding motif protein 8A			4.28
	133155	M58583	Hs.662	cerebellin 1 precursor		10.80	
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	3.00		
	133204	BE267696	Hs.254105	enkease 1, (alpha)			4.63
50	133412	U41493	Hs.73112	guanine nucleotide binding protein (G pr		12.50	
	133421	AF134160	Hs.7327	claudin 1	2.85		
	133451	AW970026	Hs.73818	ubiquitin-cytochrome c reductase hinge p			4.66
	133453	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept		6.80	
	133504	NM_004415	Hs.74316	desmoplakin (DPI, DPII)	6.14		
55	133506	BE562958	Hs.74346	hypothetical protein MGC14353			4.55
	133615	M62843	Hs.75236	ELAV (embryonic lethal, abnormal vision,			17.80
	133627	NM_002047	Hs.75280	glycyl-tRNA synthetase			4.85
	133649	U25849	Hs.75393	acid phosphatase 1, soluble			6.34
	133669	NM_006925	Hs.166975	splicing factor, arginine/serine-rich 5			14.00
60	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		6.11	
	133776	BE268649	Hs.177766	ADP-ribosyltransferase (NAD+; poly (ADP-			4.91
	133865	AB011155	Hs.170290	discs, large (Drosophila) homolog 5	3.07		
	133946	AJ001258	Hs.173878	NIPSNAP, C. elegans, homolog 1			4.60
	133973	N55540	Hs.78026	ESTs, Weakly similar to similar to ankyl			13.00
65	134047	BE262529	Hs.78771	phosphoglycerate kinase 1			3.85
	134098	BE513171	Hs.79086	mitochondrial ribosomal protein L3	2.56		
	134107	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte		8.20	
	134112	AW449809	Hs.79150	chaperonin containing TCP1, subunit 4 (d			4.08
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	31.00		
70	134160	T98152	Hs.79432	fibrillin 2 (congenital contractural era		24.60	
	134168	AA398908	Hs.181634	Homo sapiens cDNA: FLJ23602 fis, clone L			6.71
	134185	AA285136	Hs.301914	neuronal specific transcription factor D			14.74
	134201	L35035	Hs.79886	ribose 5-phosphate isomerase A (ribose 5		8.40	
	134272	X76040	Hs.278614	protease, serine, 15	4.50		
75	134276	BE083936	Hs.80976	antigen identified by monoclonal antibod		9.00	
	134353	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m			16.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	2.80		
	134380	AU077143	Hs.179565	minichromosome maintenance deficient (S.	4.68		
	134423	H53497	Hs.83006	CGI-139 protein			3.84
	134469	AA279661	Hs.83753	small nuclear ribonucleoprotein polypept			5.81
80	134470	X54942	Hs.83758	CDC28 protein kinase 2			4.21
	134498	AW246273	Hs.84131	threonyl-tRNA synthetase			7.30
	134502	BE148534	Hs.84168	UV-B repressed sequence, HUR 7		13.60	
	134510	NM_002757	Hs.250870	mitogen-activated protein kinase kinase			9.70
	134548	N95406	Hs.333495	Deleted in split-hand/split-foot 1 regio			4.63
85	134654	AK001741	Hs.8739	hypothetical protein FLJ10879	6.00		

5	134724	AF045239	Hs.321576	ring finger protein 22		12.00
	134743	AA044163	Hs.89463	potassium large conductance calcium-acti	4.00	
	134781	AA374372	Hs.89626	parathyroid hormone-like hormone		25.20
	134806	AD001528	Hs.89718	spermine synthase		4.58
	134853	BE268326	Hs.90280	5-aminimidazole-4-carboxamide ribonucle		4.79
	134859	D26488	Hs.90315	KIAA0007 protein		6.20
	134891	R51083	Hs.90787	ESTs		7.40
	134960	BE246400	Hs.285176	acetyl-Coenzyme A transporter	4.00	
	134993	BE409809	Hs.301005	purine-rich element binding protein B		4.48
10	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	9.50	
	135080	AI761180	Hs.94211	rod1 (required for cell differentiation,	5.00	
	135103	NM_003428	Hs.9450	zinc finger protein 84 (HPF2)		11.00
	135145	AW014729	Hs.95262	nuclear factor related to kappa B bindin		4.01
	135184	U13222	Hs.96028	forkhead box D1		7.00
15	135242	AI583187	Hs.9700	cyclin E1	13.50	
	135286	AW023482	Hs.97849	ESTs	6.46	
	135289	AW372569	Hs.9788	hypothetical protein MGC10924 similar to		8.80
	135355	AK001652	Hs.99423	ATP-dependent RNA helicase	10.00	
	135371	NM_008025	Hs.997	protease, serine, 22	8.00	
20	135393	L11244	Hs.99886	complement component 4-binding protein,		14.60

TABLE 5B shows the accession numbers for those primekeys lacking unigenelD's for Table 5A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
117079	1621717_1	H92325 T97125
124305	242183_1	AW963221 AA344870 AA344871 H93331
101502	18202_6	M26958
109792	754958_1	R49625 F10674
126034	1598157_1	H60340 N91637
102768	44641_1	U82321 H66077
126345	1653833_1	N49713 N49819 W03810
127066	1703458_1	R25066 R20144 R20145 Z43845
127099	244301_1	AA347668 AW956810 Z44271 F07065 F07064 R13506
119243	1774795_1	T12603 T12604
125876	1566433_1	H14480 N98295
112054	1538292_1	R43590 F10439
126979	171411_1	AA210954 AA211007
126992	880655_1	AI809521 H12174 Z42556
122318	292419_1	AA429743 AA442754
114699	135322_1	AA127386 R15644 AA127404
114793	150742_1	AA158245 AA158235
108305	111550_1	AA071391 AA069892 AA069891
108393	113411_1	AA075211 AA075245 AA075126 AA074946
100867	tigr_HT4586	U14622
123731	genbank_AA609839	AA609839
109700	genbank_F09609	F09609
120715	genbank_AA292700	AA292700
113702	genbank_T97307	T97307
115113	genbank_AA256460	AA256460
101045	entrez_J05614	J05614
108554	genbank_AA084948	AA084948
108573	genbank_AA086005	AA086005
119052	149538_1	R10889 R10888
126522	416020_1	W31912 AI167491
126605	439280_1	AA676910 AA778853 AA778865 W86800
103768	46922_1	W42667 AI560740 AI690440 AI561350 AW457906 AW151450 AI825927 AI041716 AI885600 AI742213 AW248624 AI955498 AA033947 AA845593 AI623711 N68583 C00064 AA193567 AW083868 AW163216 AA191595 AA522778 AI628008 AI915518 AA843508 AI926195 AA176265 AW167963 AA992115 W93647 AW103572 AI862994 AI342059 AA911719 AA176155 AA024712 AA069988 AA205591 AI591107 AI199673 AI811766 AI275832 AI422233 AI191852 AI096682 AI580124 AI683612 AA582453 AA927559 AA488415 T32414 AI084978 H44849 H44848 H20477 T91695 W47039 AA070055 AA024795 AA328855 AA379248 AA379330 AA385580 W25920 W03688 AA448359 AA093881 AW362477 AA089997 AI350265 W93479 N99688 AA932257 AW351469 H68590 AA663402 AA069771 AW087986 AI858420 AA600214 AI970774 AI857712 AI683081 AI855584 AW131150 AI567981 AW002714 AW189973 AW075495 AW168303 AA953714 AW516881 AI357375 AI566663 AW512676 AI570580 AI023690 AA448216 AI079853 AI422707 AA779516 AW026972 AW130082 AW162307 AW438646 AA709332 AW192394 AI167350 AI217879 AI129152 AA719509 AI350480 AA663418 AI003634 AW118546 AA180261 AA442833 AI268625 AA888881 AI038759 AAB46723 AI248770 AA993694 AI280335 AI885107 AW518649 AA641563 AA995835 AA582521 AI276744 AA436478 AI017360 AI620763 AI859887 N73926 AI076327 AI741615 AI160617 AW172819 AI492005 AA677429 AA996334 AI693771 AI950039 AI245629 AI288515 AI866186 T93293 AA173262 AA599779 AI680092 AW439316 AI084555 AI272672 AI583507 AW473219 AA738132 AW473283 AI367492 AA995410 AI689624 AA206353 AI033095 AI040382 AA873630 AI221074 AI934840 AI418680 AA844306 R94503 AA773520 AA843169 AA219425 AA629658 AI811719 AW411275 AI590981 W37907 AI591178 AI684051 AA983238 AA669347 AA976239 AA704570 AI628339 AI884391 AI241580 AI003539 AW176687 AA009650 N34566 AI333493 AI186070 AA070827 AA411683 AI280884 AA872023 AA207255 AA021576 N71953 AI885888 AW076039 T15777 AI537673 AW248048 H09554 W93480 W47001 AW079114 AA063160 AA757453 R60788 AI859431 H20478 AA218882 AA757465 AA100995 AI864135 AI934209 AA070503 H47008 AA219646 W61039 W93907 AW385050 W37957 W78028 AA189007 AA479136 R93650 AA442312 T30287 AAB47628 AA180262 AA009649 C03892 AW149464 AA310963 AA219693 AA069747 R29207 AA094784 AA293615 AA447848 AI984167 N90393 C05097 N56499 AW292351 AW149661 AW473258 AA629322 AI004409 AW105577 AI954937 AI811070 AA902422 AW514437 AA535460 AA916877 AW517122 AA974657 AA975649 AW517130 AW517129 F31737 W07688 AA193645 AA378994 AA489273 F32267 W39303 AA021181 N68810 AA406524 AA062553 AA436801 H09895 H15979 N40310

AA436789 AA232172 AW360778 W25862 R60282 AA436530 AA378894 AA187461 AI940535 AA604210 AA089514 AA360421 N88243 N84281
 AA209340 N56174 N88374 AA191088 AW247691 AA249013 AA093111 AA972536 AW298594 AA375893 T12139 W28186 AW243849
 AI288629 AA843996 W15260 AI188286 AW248079 R15836

5 119599 genbank_W45552 W45552
 112382 genbank_R59904 R59904
 105264 genbank_AA227934 AA227934
 100071 entrez_A28102 A28102
 123315 714071_1 AA496369 AA496646

10

Table 6A shows 99 genes up-regulated nonsmokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genchip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

15

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title

20

R1: average of AI for samples from non-smokers with adenocarcinoma divided by the 90th percentile of AI for samples from smokers with adenocarcinoma
 R2: average of AI for samples from non-smokers with squamous cell carcinoma divided by the 90th percentile of AI for samples from smokers with squamous cell carcinoma

Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2
100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		3.64
101174	L17330	Hs.280	pre-T/NK cell associated protein	15.00	
101296	Y12490	Hs.85092	thyroid hormone receptor interactor 11		2.46
101304	AA001021	Hs.6685	thyroid hormone receptor interactor 8		12.00
101806	AA586894	Hs.112408	S100 calcium-binding protein A7 (psoriasis)		2.68
101972	S82472		gbbeta-pol-DNA polymerase beta (axon a)		2.11
102274	U30930	Hs.158540	UDP glycosyltransferase 8 (UDP-galactose)	7.50	
102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase domain	7.50	
102832	U92015		gb:Human clone 143789 defective mariner	13.50	
103010	X52509	Hs.161640	tyrosine aminotransferase	9.50	
103439	X98266		gb:H.sapiens mRNA for ligase like protal		2.50
103563	L02911	Hs.150402	activin A receptor, type I	9.00	
103857	AK076795	Hs.45033	lacrimal proline rich protein		3.94
104239	AB002367	Hs.21355	doublecortin and CaM kinase-like 1	13.50	
104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m		12.66
104907	AA055829	Hs.196701	ESTs, Weakly similar to ALU1_HUMAN ALU	16.50	
106131	BE514788	Hs.296244	SNARE protein		2.17
106672	H47233	Hs.30643	ESTs	7.00	
106872	T56887	Hs.18282	KIAA1134 protein	11.50	
106960	AA156238	Hs.32501	ESTs		2.38
106971	Z43846	Hs.194478	Homo sapiens mRNA; cDNA DKFZp434O1572 (f	9.50	
107982	AA035375	Hs.57887	ESTs, Weakly similar to KIAA0758 protal		2.95
108562	AA100796		gb:zm26c06.s1 Stratagene pancreas (93720	16.50	
108599	AB018549	Hs.69328	MD-2 protein	13.00	
108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypotheti		2.40
109247	AA314907	Hs.85950	ESTs	7.00	
109630	R44607	Hs.22672	ESTs		5.00
110193	AI004874	Hs.310764	Homo sapiens mRNA; cDNA DKFZp434M082 (fr	12.50	
110234	H24458	Hs.32085	EST	16.50	
110644	R94207	Hs.268989	ESTs, Highly similar to type II CALM/AF1	8.00	
110886	AW274992	Hs.72249	three-PDZ containing protein similar to	17.00	
111057	T79639	Hs.14629	ESTs	16.50	
111950	AF071594	Hs.110457	Wolf-Hirschhorn syndrome candidate 1	11.00	
112291	R53972	Hs.26026	ESTs		3.00
112956	Z43784	Hs.75893	ankyrin 3, node of Ranvier (ankyrin G)		2.79
113009	T23699	Hs.7246	ESTs		4.50
113060	BE564162	Hs.250820	hypothetical protein FLJ14827	9.79	
113073	N39342	Hs.103042	microtubule-associated protein 1B	32.50	
113074	AK001335	Hs.31137	protein tyrosine phosphatase, receptor t		3.82
113121	T48011	Hs.8764	EST		2.21
113125	AA968672	Hs.8929	hypothetical protein FLJ11362	19.50	
113757	AA703095	Hs.18631	ESTs		2.65
113848	W52854	Hs.27099	hypothetical protein FLJ23293 similar to	6.00	
113884	AI333076	Hs.28529	chromosome 12 open reading frame 2		6.00
113936	W17056	Hs.83623	nuclear receptor subfamily 1, group I, m		4.63
114875	AA235609	Hs.236443	Homo sapiens mRNA; cDNA DKFZp564N1063 (7.00
114987	AA251016	Hs.87808	EST		6.00
115460	AW958439	Hs.38613	ESTs		2.27
115722	W91892	Hs.59609	ESTs		9.00
116261	AA481788	Hs.190150	ESTs	9.50	
116830	H61037	Hs.70404	ESTs, Weakly similar to ALU2_HUMAN ALU	8.50	
116970	AB023179	Hs.9059	KIAA0962 protein	7.50	
117178	H98675	Hs.269034	ESTs		2.68
117757	AF088019	Hs.46732	EST	7.50	
118283	AA287747	Hs.173012	ESTs, Weakly similar to A46010 X-linked	16.50	
118384	AF217525	Hs.49002	Down syndrome cell adhesion molecule		2.50
118657	AI822106	Hs.49902	ESTs		2.39
120328	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapi		3.50
120404	AB023230	Hs.96427	KIAA1013 protein	7.00	
120524	AA261852	Hs.192905	ESTs	6.00	
120688	AW207555	Hs.97093	Homo sapiens cDNA: FLJ23004 fts, clone L	17.92	

5	121558	AA412497	gb:z195g12.s1 Soares_testis_NHT Homo sap		2.95
	121676	H56037	ESTs	10.00	
	121936	AI024600	ESTs	15.00	
	121938	AA428659	ESTs	14.00	
	122177	AA435789	EST	8.93	
10	123442	AA289652	Hs.111496 Homo sapiens cDNA FLJ11643 fis, clone HE	13.04	
	123551	AA608837	gb:af03h12.s1 Soares_testis_NHT Homo sap	11.50	
	123756	AA609971	EST	11.00	
	123861	AA620840	gb:af09g01.s1 Soares_testis_NHT Homo sap		2.50
	124371	N24924	ESTs	6.50	
15	127477	BE328720	ESTs		4.33
	127591	AI190540	ESTs		3.02
	128252	AA455924	ESTs	7.00	
	128426	AI265784	ESTs		2.08
	128925	R67419	Hs.21851 Homo sapiens cDNA FLJ12900 fis, clone NT		2.11
20	128945	AI990506	Hs.8077 Homo sapiens mRNA; cDNA DKFZp547E184 (fr	10.00	
	129105	AI769160	Hs.108681 Homo sapiens brain tumor associated prot	15.50	
	129235	AW577238	Hs.126084 KIAA1055 protein		4.25
	129506	AB020684	Hs.11217 KIAA0877 protein	6.50	
	129595	U09550	Hs.1154 oviductal glycoprotein 1, 120kD (mucin 9		10.00
25	130160	AA305688	Hs.267695 UDP-Gal:betaGlcNAc beta 1,3-galactosyltr	20.00	
	130340	D82326	Hs.239106 solute carrier family 3 (cystine, dibasi	11.50	
	131220	AB023194	Hs.300855 KIAA0977 protein	17.50	
	131430	AI879148	Hs.26770 fatty acid binding protein 7, brain	6.10	
	132114	NM_006152	Hs.40202 lymphoid-restricted membrane protein		6.15
30	132458	AA935315	Hs.48965 Homo sapiens cDNA: FLJ21693 fis, clone C		5.58
	132647	NM_006927	Hs.54432 sialyltransferase 4B (beta-galactosidase	7.50	
	132655	D49372	Hs.54460 small inducible cytokine subfamily A (Cy		2.53
	132682	AI077500	Hs.54900 serologically defined colon cancer antig		2.50
	132747	AA345241	Hs.55950 ESTs, Weakly similar to KIAA1330 protein		2.83
35	132812	R50333	Hs.92186 Leman coiled-coil protein		3.82
	133337	AF085983	Hs.293676 ESTs		5.00
	133876	AL134906	Hs.771 phosphorylase, glycogen; liver (Hers dis		3.00
	134119	AW157837	Hs.79226 fasciculation and elongation protein zel		2.06
	134464	AA302983	Hs.239720 CCR4-NOT transcription complex, subunit		2.27
40	134542	M14156	Hs.85112 insulin-like growth factor 1 (somatomedi		11.50
	135002	AA448542	Hs.251677 G antigen 7B	87.00	
	135305	AA203555	Hs.98288 Homo sapiens cDNA FLJ14903 fis, clone PL		6.50

TABLE 6B show the accession numbers for those primekeys lacking unigenes/Ds for Table 6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
108562	36375_1	AA100796 AF020589 AA074629 AA075946 AA100849 AA085347 AA126309 AA079311 AA079323 AA085274
103439	35330_1	X98266 N41124
123551	genbank_AA608837	AA608837
123861	genbank_AA620840	AA620840
102832	entrez_U92015	U92015
101972	entrez_S82472	S82472
121558	genbank_AA412497	AA412497

Table 7A shows 98 genes down-regulated in non-smokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
10	R1:	90th percentile of AI for samples from smokers with adenocarcinoma divided by the average of AI for samples from non-smokers with adenocarcinoma.			
	R2:	90th percentile of AI for samples from smokers with squamous cell carcinoma divided by the average of AI for samples from non-smokers with squamous cell carcinoma.			
	Pkey	ExAccn	UnigeneID	Unigene Title	R1 R2
15	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	164.10
	100380	D82343	Hs.18551	neuroblastoma (nerve tissue) protein	77.40
	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptide	102.40
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte	463.80
	101046	K01160		(NONE)	672.00
20	101056	AW970254	Hs.889	Charot-Leyden crystal protein	66.00
	101175	U82671	Hs.36980	melanoma antigen, family A, 2	77.20
	101497	W05150	Hs.37034	homeo box A5	62.80
	101663	NM_003528	Hs.2178	H2B histone family, member Q	78.00
	101677	NM_000715	Hs.1012	complement component 4-binding protein,	186.20
25	101745	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	80.08
	101941	S77583		gb:HERVK10/HUMMTV reverse transcriptase	99.20
	102125	NM_006456	Hs.288215	slalyltransferase	103.10
	102242	U27185	Hs.82547	retinoic acid receptor responder (lazarus	67.00
	102340	U37055	Hs.278657	macrophage stimulating 1 (hepatocyte gro	71.60
30	102359	U39840	Hs.299867	hepatocyte nuclear factor 3, alpha	69.70
	102457	NM_001394	Hs.2359	dual specificity phosphatase 4	153.00
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	65.70
	102796	AL079646	Hs.107019	sympleskin; Huntingtin interacting protei	58.80
	102829	NM_006183	Hs.80962	neurotensin	258.80
35	103207	X72790		gb:Human endogenous retrovirus mRNA for	70.00
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	212.10
	103260	X78416	Hs.3155	casein, alpha	130.70
	103351	X89211		gb:H.sapiens DNA for endogenous retrovir	64.60
40	104212	AB002298	Hs.173035	KIAA0300 protein	66.80
	104252	AF002246	Hs.210863	cell adhesion molecule with homology to	63.80
	104258	AF007216	Hs.5462	solute carrier family 4, sodium bicarbon	94.40
	105024	AA126311	Hs.9879	ESTs	68.20
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S	74.60
45	106440	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	71.10
	106566	BE288210		gb:601118016F1 NIH_MGC_17 Homo sapiens c	73.20
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	83.80
	106614	AA648459	Hs.335951	hypothetical protein AF301222	62.30
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	202.40
	106999	H93281	Hs.10710	hypothetical protein FLJ20417	89.60
50	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c	66.40
	108810	AW295647	Hs.71331	hypothetical protein MGCS350	85.50
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	63.40
	109597	AA989362	Hs.293780	ESTs	85.00
55	109691	T65568	Hs.12860	ESTs	58.70
	109704	AI743880	Hs.12876	ESTs	60.60
	110942	R63503	Hs.28419	ESTs	76.40
	111722	R23924	Hs.23596	EST	74.60
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-ii	64.80
60	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	76.70
	113073	N39342	Hs.103042	microtubule-associated protein 18	120.20
	114251	H15261	Hs.21948	ESTs	127.20
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	174.00
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	91.00
65	115815	AW905328	Hs.180842	ribosomal protein L13	66.40
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	226.60
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970	82.80
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	361.60
	116552	D20508	Hs.164649	hypothetical protein DKFZp434H247	69.00
70	116571	D45652		gb:HUMGS02848 Human adult lung 3' direct	64.20
	118466	N66741		gb:yz33g08.s1 Morton Fetal Cochlea Homo	63.50
	120484	AA253170	Hs.96473	EST	81.60
	120983	AA398209	Hs.97587	EST	81.10
	121034	AL389951	Hs.271623	nucleoporin 50kD	66.20
75	121423	AW973352	Hs.290585	ESTs	64.40
	122553	AA451884	Hs.190121	ESTs	60.40
	122946	AI718702	Hs.308026	major histocompatibility complex, class	188.60
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H	80.20
	124472	N52517	Hs.102670	EST	71.00
80	124526	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	104.90
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A	72.00
	125731	R61771	Hs.26912	ESTs	69.90
	125747	NM_002884	Hs.865	RAP1A, member of RAS oncogene family	69.00
	126020	H79863	Hs.114243	ESTs	62.40
	126547	U47732	Hs.84072	transmembrane 4 superfamily member 3	62.80
85	126966	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra	60.10

	127472	AA761378	Hs.192013	ESTs	70.20	
	127610	AA950867	Hs.150271	ESTs, Highly similar to unnamed protein	64.00	
	127742	AW293496	Hs.180138	ESTs	85.20	
	127967	AI022103	Hs.124511	ESTs	96.60	
5	128233	AW889132	Hs.11916	ribokinase		78.90
	128420	AA650274	Hs.41296	fibronectin leucine rich transmembrane p		105.90
	128766	AW160432	Hs.296460	craniofacial development protein 1	66.80	
	129014	AW935187	Hs.170162	KIAA1357 protein		58.53
	129215	AB040930	Hs.126085	KIAA1497 protein	64.20	
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 18)	63.80	
	130385	AW067800	Hs.155223	stannocalcin 2		139.60
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)		64.60
	131025	AB040900	Hs.6189	KIAA1467 protein	64.40	
	131241	BE501914	Hs.24654	Homo sapiens cDNA FLJ11640 fis, clone HE	76.20	
15	131775	AB014548	Hs.31921	KIAA0648 protein	97.80	
	132240	AB018324	Hs.42676	KIAA0781 protein		71.00
	132856	NM_001448	Hs.58367	glypican 4		88.40
	132977	AA093322	Hs.301404	RNA binding motif protein 3	133.20	
	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		59.30
20	133818	AI110684	Hs.7645	fibrinogen, B beta polypeptide	341.00	
	134264	AF149297	Hs.8087	NAG-5 protein		64.30
	134265	M83772	Hs.80876	flavin containing monooxygenase 3		232.53
	134346	X84002	Hs.82037	TATA box binding protein (TBP)-associate	66.00	
	134395	AA456539	Hs.8262	lysosomal-associated membrane protein 2		75.80
25	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su		108.30
	135056	N75765	Hs.93765	lipoma HMGIC fusion partner	71.40	
	135309	AI584123	Hs.42500	ADP-ribosylation factor-like 5	70.40	

30 TABLE 7B shows the accession numbers for those primekeys lacking unigenelD's for Table 7A. For each probaset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

35 Pkey: Unique Eos probaset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
103207	30635_4	X72790
106566	120358_1	BE298210 AI672315 AW086489 BE298417 AA455921 AA902537 BE327124 R14963 AA085210 AW274273 AI333584 AI369742 AI039658 AI885095 AI476470 AI287650 AI885299 AI985381 AW592624 AW340135 AI266556 AA456390 AI310815 AA484951
116571	genbank_D45652	D45652
118466	genbank_N66741	N66741
101046	entrez_K01160 K01160	
101941	entrez_S77583 S77583	
103351	entrez_X89211 X89211	
123130	genbank_AA487200	AA487200

Table 8A shows 1720 genes either up or down-regulated in lung tumors or chronically diseased lung relative to a broad collection of over 40 distinct normal body tissues. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 39494 probesets on the Eos/Affymetrix Hu02 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5

Pkey: Unique Eos probeset identifier number
ExAccn: Exemplar Accession number, Genbank accession number
UnigeneID: Unigene number
Unigene Title: Unigene gene title

10

R1: 70th percentile of AI for lung tumors divided by 90th percentile of AI for normal lung
R2: 70th percentile of AI for chronically diseased lung divided by 90th percentile of AI for normal lung

	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
15	300097	AJ916973	Hs.213603	ESTs	5.46	4.69
	300117	AW189787	Hs.147474	ESTs	0.58	0.56
	300197	AJ686661	Hs.218286	ESTs	4.26	5.44
	300201	AJ308300		gb:ta90c06.x1 NCI_CGAP_Bm20 Homo sapien	0.62	0.83
	300225	AJ989963	Hs.197505	ESTs	1.68	1.75
20	300247	AW274682	Hs.161394	ESTs	1.08	2.28
	300256	AJ469095	Hs.298241	Transmembrane protease, serine 3	0.86	1.00
	300337	AJ707881	Hs.202090	ESTs	5.80	9.09
	300362	Z42308		gb:HSC0FB121 normalized infant brain cDN	4.18	12.78
	300374	AJ859947	Hs.314158	ESTs	2.99	4.38
25	300387	AW270150	Hs.254516	ESTs	1.50	2.53
	300440	AJ421541	Hs.146164	ESTs	3.98	5.25
	300441	R10367	Hs.307921	EST, Weakly similar to Z232_HUMAN ZINC F	3.18	6.80
	300449	AJ362967	Hs.132221	hypothetical protein FLJ12401	0.43	0.62
	300469	AW135830	Hs.233955	hypothetical protein FLJ20401	0.16	0.83
30	300552	K85711	Hs.21838	hypothetical protein FLJ11191	4.10	9.75
	300627	W27363		gb:ab37d01.r1 Stratagene HeLa cell s3 93	4.60	12.60
	300630	AW118822	Hs.128757	ESTs	2.91	5.85
	300716	AJ216113	Hs.126280	hypothetical protein FLJ23393	1.00	0.92
	300738	AJ623332	Hs.130541	KIAA1542 protein	1.82	1.71
35	300777	AA235361	Hs.96840	KIAA1527 protein	4.48	8.22
	300790	AJ492471	Hs.188270	ESTs	1.29	1.18
	300832	AJ588147	Hs.220515	ESTs, Weakly similar to T03829 transcrip	5.51	8.56
	300836	Z44942	Hs.22958	calcium channel alpha2-delta3 subunit	4.90	6.34
	300838	AJ582897	Hs.192570	hypothetical protein FLJ22028	1.70	2.81
40	300878	AW449802	Hs.285901	Homo sapiens cDNA FLJ20428 fis, clone KA	4.56	7.91
	300897	AJ890356	Hs.127804	ESTs, Weakly similar to T17233 hypotheti	2.23	1.58
	300926	AA504860		gb:ab03a10.s1 Stratagene fetal retina 93	2.13	3.50
	300950	AJ041019	Hs.152454	ESTs	2.74	4.46
	300961	AW204069	Hs.312716	ESTs, Weakly similar to unnamed protein	1.00	1.00
45	300962	AA593373	Hs.293744	ESTs	1.46	1.51
	300967	AA565209	Hs.269439	ESTs	0.39	1.30
	300987	AW450840	Hs.148590	ESTs, Weakly similar to AF208846 1 BM-00	1.49	1.08
	300988	AJ927208	Hs.208952	ESTs	0.16	0.37
	301050	AW136973	Hs.288516	ESTs, Weakly similar to S69890 mitogen i	3.23	1.94
50	301098	AA677570	Hs.185918	ESTs	6.76	14.28
	301157	AA729905	Hs.231916	ESTs	3.16	8.85
	301162	AJ142118	Hs.129004	ESTs	1.68	7.18
	301170	AA737594	Hs.247606	ESTs	4.40	6.42
	301192	AJ808751	Hs.121188	ESTs	6.38	11.59
55	301193	AA758115	Hs.128350	ESTs, Weakly similar to JC5423 2-hydroxy	4.35	7.78
	301267	AW297762	Hs.255690	ESTs	1.66	1.61
	301281	AA843965	Hs.190586	ESTs	2.19	1.78
	301341	AJ819198	Hs.208229	ESTs	0.76	0.76
	301382	AA912839	Hs.163369	ESTs	1.00	1.81
60	301407	AW450466	Hs.126830	ESTs	1.48	1.51
	301452	AA975688	Hs.159955	ESTs	0.51	1.46
	301483	AW272467	Hs.254655	Untitled	2.40	5.02
	301494	AJ678034	Hs.131099	ESTs	2.79	3.41
	301521	AJ733621	Hs.133011	zinc finger protein 117 (HPF9)	0.67	0.67
65	301531	AJ077462	Hs.134084	ESTs	2.52	3.76
	301580	AJ878959	Hs.73737	splicing factor, arginine/serine-rich 1	7.41	11.92
	301676	Z43570	Hs.27453	ESTs, Moderately similar to G01251 Rar p	8.31	10.70
	301690	F05865	Hs.108323	ubiquitin-conjugating enzyme E2E 2 (homo	2.70	4.22
	301718	F07744	Hs.7987	DKFZP434F162 protein	4.20	8.78
70	301799	AA364252	Hs.286132	D15F37 (pseudogene)	5.93	7.04
	301804	AA581004	Hs.62180	anillin (Drosophila Scraps homolog), act	1.70	0.76
	301822	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	1.58	1.36
	301846	R20002	Hs.6823	hypothetical protein FLJ10430	1.00	1.00
	301868	T71508	Hs.13861	ESTs, Weakly similar to pH sensitive max	2.88	5.49
75	301882	T78054		gb:yc97g09.r1 Soares infant brain 1N1B H	2.28	3.80
	301905	AJ991127	Hs.117202	ESTs	1.00	1.00
	301948	AA344647	Hs.116724	aldo-keto reductase family 1, member B11	5.28	2.28
	301960	AW070252	Hs.27973	KIAA0874 protein	5.38	6.48
	302011	T91418	Hs.125156	transcriptional adaptor 2 (ADA2, yeast,	3.03	3.42
80	302016	N40834	Hs.23495	hypothetical protein FLJ11252	1.00	1.25
	302041	NM_001501	Hs.129715	gonadotropin-releasing hormone 2	0.71	0.99
	302072	AJ238381	Hs.132576	paired box gene 9	1.60	1.71
	302094	AJ286176	Hs.6786	ESTs	0.52	1.20
	302095	AW044300	Hs.137506	Homo sapiens BAC clone RP11-120J2 from 7	2.75	4.93
85	302148	AW269618	Hs.23244	ESTs	3.04	3.87

	302155	AI088485	Hs.144759	ESTs	0.45	1.15
	302201	AJ006276	Hs.159003	transient receptor potential channel 6	0.33	0.84
	302202	AF097159	Hs.159140	UDP-Gal4betaGlcNAc beta 1,4- galactosylt	0.52	0.94
5	302206	AI937193	Hs.41143	phosphoinositide-specific phospholipase	2.76	3.65
	302209	AF047445	Hs.159297	killer cell lectin-like receptor subfam1	1.00	1.00
	302235	AL049987	Hs.166361	Homo sapiens mRNA; cDNA DKFZp564F112 (fr	1.68	1.50
	302290	AL117607	Hs.175563	Homo sapiens mRNA; cDNA DKFZp564N0763 (f	1.00	2.11
	302328	AA354849	Hs.23240	Homo sapiens cDNA FLJ13496 fis, clone PL	9.38	13.08
10	302346	AL039101	Hs.194625	dynein, cytoplasmic, light intermediate	3.27	7.24
	302360	AJ010901	Hs.198267	mucin 4, tracheobronchial	2.54	1.88
	302384	Y08982	Hs.202676	synaptonemal complex protein 2	1.00	0.91
	302406	U86751	Hs.211956	CD3-epsilon-associated protein; antisens	2.63	2.67
	302409	AF155156	Hs.218028	adaptor-related protein complex 4, epsil	5.82	9.34
15	302423	AB028977	Hs.225974	KIAA1054 protein	3.66	3.18
	302432	AL080068	Hs.272534	Homo sapiens mRNA; cDNA DKFZp564J062 (fr	2.44	6.77
	302435	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	0.44	0.84
	302437	AB024730	Hs.227473	UDP-N-acetylglucosaminase-1,3-D-mannosid	4.18	5.64
	302455	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	1.85	0.92
20	302472	AA317451	Hs.6335	SWI/SNF related, matrix associated, acti	2.04	2.13
	302476	AF182294	Hs.241578	U6 snRNA-associated Sm-like protein LSM8	1.44	1.89
	302489	T80660	Hs.230424	Homo sapiens cDNA FLJ13540 fis, clone PL	0.51	1.10
	302490	AA885502	Hs.187032	ESTs	2.64	4.87
	302562	AJ005585	Hs.48956	gap junction protein, beta 6 (connexin 3	5.34	2.68
25	302566	AA085996	Hs.248572	hypothetical protein FLJ22965	1.00	1.21
	302630	AB029488	Hs.272100	SMS3 protein	0.52	1.24
	302634	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	1.00	1.00
	302638	AA463798	Hs.102696	MCT-1 protein	1.58	1.02
	302647	X57723	Hs.198273	NADH dehydrogenase (ubiquinone) 1 beta s	2.72	6.85
30	302655	AJ227892	Hs.146274	ESTs	1.00	4.32
	302656	AW293005	Hs.70704	Homo sapiens, clone IMAGE:2823731, mRNA,	2.97	0.93
	302668	AA580691	Hs.180789	S164 protein	0.80	0.95
	302679	H65022		gb:yu66g11.1 r1 Weizmann Olfactory Epithel	1.68	5.04
	302680	AW192334	Hs.38218	ESTs	2.70	7.98
35	302697	AJ001408		gb:Homo sapiens mRNA for immunoglobulin	4.25	8.13
	302705	U09060		gb:Human immunoglobulin heavy chain, V-r	3.91	8.68
	302711	L08442		gb:Human autonomously replicating sequen	2.20	2.73
	302719	W69724	Hs.288959	hypothetical protein FLJ20920	0.54	1.02
	302742	L12069		gb:Homo sapiens (clone WR4.10VH) anti-th	4.28	11.57
40	302755	AW384815	Hs.149208	KIAA1555 protein	1.57	2.38
	302771	H98478	Hs.42522	ESTs	2.94	4.68
	302789	AJ245067		gb:Homo sapiens mRNA for immunoglobulin	3.49	6.31
	302795	AJ245313	Hs.272838	hypothetical protein FLJ10494	0.80	2.74
	302802	Y08250		gb:H.sapiens mRNA for variable region of	1.13	0.77
45	302803	AA442824	Hs.293961	ESTs, Moderately similar to putative DNA	3.14	10.68
	302812	N31301	Hs.152664	hypothetical protein FLJ20051	3.04	8.24
	302847	X98940		gb:H.sapiens rearranged Ig heavy chain (1.80	1.92
	302885	AL137763	Hs.132127	hypothetical protein LOC57822	1.00	1.00
	302943	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	0.53	0.67
50	302977	AW263124	Hs.315111	hypothetical protein FLJ12894	2.45	2.62
	303006	AF078950	Hs.24139	Homo sapiens cDNA: FLJ23137 fis, clone L	4.88	8.61
	303011	AF090405		gb:Homo sapiens clone 2A1 scFV antibody	1.41	1.86
	303013	F07898	Hs.288968	RAB22A, member RAS oncogene family	1.51	1.19
	303061	AF151882	Hs.27693	peptidyl/prolyl isomerase (cyclophilin)-I	0.72	0.76
55	303077	AF163305		gb:H.sapiens T-cell receptor mRNA	1.17	3.90
	303090	AA442559	Hs.146286	kinesin family member 13A	4.08	6.46
	303091	AF192913	Hs.130683	zinc finger protein 180 (HHz168)	2.50	4.37
	303094	AF195513	Hs.278953	Pur-gamma	5.38	8.38
	303095	AF202051	Hs.134079	NM23-H8	3.26	4.08
60	303131	AW081061	Hs.103180	DC2 protein	2.02	1.83
	303195	AA082211	Hs.233936	myosin, light polypeptide, regulatory, n	1.32	3.95
	303196	AA082298	Hs.59710	ESTs	0.77	0.53
	303216	AA581439	Hs.152328	ESTs	0.24	0.63
	303222	AA333538	Hs.204501	hypothetical protein FLJ10534	3.56	6.22
65	303234	AA132255	Hs.143951	ESTs	2.28	3.17
	303251	AW340037	Hs.115897	protocadherin 12	0.38	1.02
	303295	AA205625	Hs.208067	ESTs	2.30	1.00
	303297	T80072	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.86	4.48
	303316	AF033122	Hs.141125	p53 regulated PA26 nuclear protein	0.10	0.80
70	303467	AA398801	Hs.323397	ESTs	4.54	9.65
	303506	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	0.09	0.04
	303552	AA359799	Hs.224662	ESTs, Weakly similar to unnamed protein	1.00	1.72
	303598	AA382814		gb:EST96097 Testis I Homo sapiens cDNA 5	4.96	9.14
	303637	AF056083	Hs.24879	phosphatidic acid phosphatase type 2C	2.05	2.02
75	303655	AA504702	Hs.258802	ATPase, (Na+)/K+ transporting, beta 4 po	1.00	1.24
	303756	AI738488	Hs.115838	ESTs	1.08	1.43
	303856	AA958589	Hs.180532	glucose phosphate isomerase	1.76	1.31
	303893	N88597	Hs.113503	karyopherin (importin) beta 3	2.30	2.57
	303907	AW467774	Hs.171880	polymerase (RNA) II (DNA directed) polyp	3.10	5.79
80	303946	AW474195	Hs.306637	Homo sapiens cDNA FLJ12363 fis, clone MA	5.06	11.86
	303978	AW513315		gb:xa043c12x1 NCL_CGAP_UH1 Homo sapiens	5.14	7.31
	303981	AW513804	Hs.278834	ESTs, Weakly similar to ALU1_HUMAN ALU S	2.83	4.06
	303990	AW515465		gb:cu71a11x1 NCL_CGAP_Kid8 Homo sapiens	1.15	2.35
	303998	AW516449		gb:xi58f05x1 NCL_CGAP_U12 Homo sapiens	2.20	9.35
85	303999	AW516611		gb:cp70b11x1 NCL_CGAP_Ov39 Homo sapiens	4.85	6.28
	304006	AW517947		gb:xt65h02x1 NCL_CGAP_U12 Homo sapiens	3.21	4.07

5	304008	AW518198	Hs.3297	ribosomal protein S27a	6.50	11.08
	304009	AW518206	Hs.181165	eukaryotic translation elongation factor	1.88	3.27
	304024	T03036		gb:FB21B7 Fetal brain, Stratagene Homo s	2.15	3.55
	304026	T03160		gb:FB26F2 Fetal brain, Stratagene Homo s	5.88	11.80
	304028	T03266		gb:FB7C1 Fetal brain, Stratagene Homo sa	5.59	13.46
10	304036	T16855	Hs.244621	ribosomal protein S14	6.55	14.43
	304046	T54803		gb:yb42d06.s1 Stratagene fetal spleen (9	6.18	12.19
	304061	T61521		gb:ycb73g01.s1 Stratagene ovary (937217)	2.64	8.23
	304063	T62536		gb:ycd04c12.s1 Stratagene lung (937210) H	0.53	1.61
	304097	R25376	Hs.177592	ribosomal protein, large, P1	6.49	11.67
15	304114	R78946		gb:yl87g02.s1 Soares placenta Nb2HP Homo	2.90	4.18
	304122	H28966		gb:ym31a06.s1 Soares infant brain 1N1B H	1.00	2.76
	304155	H68696		gb:yr78b06.s1 Soares fetal liver spleen	0.79	1.18
	304203	N56929		gb:yy82d08.s1 Soares_multiple_sclerosis_	4.28	11.34
	304234	W81608		gb:zd98h06.s1 Soares_fetal_heart_NbHH19W	6.47	11.03
20	304267	AA064862	Hs.73742	ribosomal protein, large, P0	1.34	1.16
	304270	AA069711	Hs.297753	vimentin	3.40	5.40
	304287	AA079286	Hs.78466	proteasome (prosome, macropain) 26S sub	2.93	4.42
	304348	AA179968		gb:zp38g12.s1 Stratagene muscle 937209 H	3.98	10.96
	304415	AA290747	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	3.32	5.99
25	304430	AA347682		gb:EST54044 Fetal heart II Homo sapiens	1.00	1.00
	304456	AA411240		gb:zy25g05.s1 Soares_NhHMPu_S1 Homo sapi	1.42	3.33
	304521	AA464716		gb:zx82c11.s1 Soares ovary tumor NbHOT H	2.18	1.15
	304526	AA476427		gb:zx02c05.s1 Soares_total_fetus_Nb2HF8_	5.38	14.11
	304542	AA482602	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.16	8.23
30	304546	AA486074	Hs.297681	serine (or cysteine) proteinase inhibitor	0.55	1.20
	304607	AA513322		gb:nh85e08.s1 NCI_CGAP_Br1.1 Homo sapien	1.95	2.10
	304640	AA524440	Hs.111334	ferritin, light polypeptide	2.10	2.83
	304650	AA527489	Hs.3463	ribosomal protein S23	3.33	12.62
	304735	AA576453		gb:nm75h11.s1 NCI_CGAP_Co9 Homo sapiens	1.33	0.88
35	304760	AA580401		gb:nn13g09.s1 NCI_CGAP_Co12 Homo sapiens	3.68	8.14
	304849	AA588157	Hs.13801	KIAA1685 protein	2.77	3.70
	304917	AA602585	Hs.284136	PRO2047 protein	7.16	11.01
	304921	AA603092	Hs.297753	vimentin	2.47	4.24
	304966	AA613893	Hs.282435	ESTs	6.78	11.66
40	304987	AA618044	Hs.300697	immunoglobulin heavy constant gamma 3 (G	0.90	1.23
	305016	AA626876		gb:zu89h06.s1 Soares_testis_NHT Homo sap	6.46	10.17
	305034	AA630128		gb:ab99c04.s1 Stratagene lung (937210) H	1.00	1.00
	305072	AA641012		gb:nr72a12.s1 NCI_CGAP_Pr24 Homo sapiens	5.68	11.59
	305111	AA644187	Hs.303405	ESTs	1.48	1.37
45	305148	AA654070		gb:nt01g08.s1 NCI_CGAP_Lym3 Homo sapiens	1.76	4.61
	305159	AA659166	Hs.275668	EST, Weakly similar to EF1D_HUMAN ELONG	1.00	2.15
	305190	AA665955		gb:ag57d12.s1 Gessler Wilms tumor Homo s	5.31	8.14
	305232	AA670052	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	0.78	1.18
	305235	AA670480		gb:ag37e01.s1 Jla bone marrow stroma Hom	3.11	8.66
50	305245	AA676695	Hs.81328	nuclear factor of kappa light polypeptid	4.38	7.53
	305312	AA700201		gb:zj44d07.s1 Soares_fetal_liver_spleen_	2.13	2.66
	305322	AA701597	Hs.163019	EST	1.20	1.40
	305394	AA720942	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.16	0.68
	305413	AA724659		gb:al10f08.s1 Soares_parathyroid_tumor_N	5.86	9.87
55	305447	AA737856		gb:mx10c08.s1 NCI_CGAP_GC3 Homo sapiens	2.21	2.86
	305476	AA745664	Hs.287445	hypothetical protein FLJ11726	3.36	6.54
	305483	AA748030	Hs.303512	EST	1.00	2.02
	305528	AA769156		gb:nz12e05.s1 NCI_CGAP_GCB1 Homo sapiens	6.44	9.10
	305612	AA782347	Hs.272572	hemoglobin, alpha 2	0.19	0.79
60	305614	AA782866		gb:aj09h02.s1 Soares_parathyroid_tumor_N	1.00	1.00
	305616	AA782884	Hs.275865	ribosomal protein S18	7.57	10.20
	305637	AA806124		gb:oe29a12.s1 NCI_CGAP_Pr25 Homo sapiens	4.78	12.42
	305639	AA806138		gb:oe29c12.s1 NCI_CGAP_Pr25 Homo sapiens	0.89	0.70
	305650	AA807709		gb:nw31e04.s1 NCI_CGAP_GC80 Homo sapiens	4.49	8.71
65	305690	AA813477		gb:ai67a05.s1 Soares_testis_NHT Homo sap	4.91	9.40
	305726	AA828156	Hs.73742	ribosomal protein, large, P0	0.19	0.81
	305728	AA828209		gb:of34a02.s1 NCI_CGAP_Kid6 Homo sapiens	5.12	9.29
	305759	AA835353		gb:ak72b06.s1 Barstead spleen HPLRB2 Hom	1.66	4.11
	305792	AA845256		gb:ak84a08.s1 Barstead spleen HPLRB2 Hom	2.34	4.25
70	305864	AA864374	Hs.73742	ribosomal protein, large, P0	0.30	1.40
	305901	AA872958		gb:oh63h08.s1 NCI_CGAP_Kid5 Homo sapiens	2.10	5.21
	305910	AA875981		gb:mx21h02.s1 NCI_CGAP_GC3 Homo sapiens	0.32	1.01
	306015	AA897116		gb:am08b07.s1 Soares_NFL_T_GBC_S1 Homo s1.56	1.65	1.12
	306017	AA897221	Hs.109058	ribosomal protein S6 kinase, 90kD, polyp	5.21	7.90
75	306020	AA897630	Hs.130027	EST	1.96	6.59
	306063	AA906316		gb:ok03g03.s1 Soares_NFL_T_GBC_S1 Homo s	7.38	20.69
	306065	AA906725		gb:ok78g02.s1 NCI_CGAP_GC4 Homo sapiens	7.19	13.48
	306104	AA910956		gb:ok85h11.s1 NCI_CGAP_Kid3 Homo sapiens	6.50	9.13
	306109	AA911861		gb:og21a07.s1 NCI_CGAP_PNS1 Homo sapiens	4.21	5.25
80	306148	AA917409	Hs.288036	tRNA isopentenylpyrophosphate transferas	2.20	2.70
	306242	AA932805		gb:oo60g04.s1 NCI_CGAP_Lu5 Homo sapiens	2.84	5.35
	306288	AA936900		gb:ol53h05.s1 NCI_CGAP_HN3 Homo sapiens	1.60	1.12
	306325	AA953072	Hs.210546	interleukin 21 receptor	1.65	2.26
	306353	AA961382	Hs.275865	ribosomal protein S18	3.78	6.32
85	306375	AA968650	Hs.276018	EST, Moderately similar to JC4662 ribos	4.30	5.74
	306396	AA970223		gb:op09d05.s1 NCI_CGAP_Kid6 Homo sapiens	0.95	2.45
	306428	AA975110	Hs.191228	hypothetical protein FLJ20284	3.19	4.10
	306442	AA976899		gb:og35e09.s1 NCI_CGAP_GC4 Homo sapiens	4.67	7.44
	306446	AA977348		gb:og72e12.s1 NCI_CGAP_Kid6 Homo sapiens	3.92	6.27

5	306458	AA978186		gb:op33c06.s1 Soares_NFL_T_GBC_S1 Homo s	3.35	5.77
	306467	AA983508	Hs.163593	ribosomal protein L18a	3.72	5.37
	306510	AA988546		gb:or84d07.s1 NCL_CGAP_Lu5 Homo sapiens	1.00	1.00
	306555	AA994304	Hs.276083	EST, Weakly similar to RL23_HUMAN 60S R	6.61	10.91
	306557	AA994530		gb:ou57e08.s1 NCL_CGAP_Br2 Homo sapiens	16.20	31.83
10	306572	AA995686		gb:os25c12.s1 NCL_CGAP_Kid5 Homo sapiens	2.51	6.52
	306582	AA996248		gb:os18c10.s1 NCL_CGAP_Kid5 Homo sapiens	1.42	3.13
	306598	AI000320	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.91	8.68
	306605	AI000497	Hs.119500	ribosomal protein, large P2	1.96	8.60
	306655	AI004024		gb:ou11b07.x1 Soares_NFL_T_GBC_S1 Homo s	0.11	0.45
15	306676	AI005603	Hs.284136	PRO2047 protein	9.56	17.28
	306686	AI015615		gb:ov29f10.x1 Soares_testis_NHT Homo sap	1.86	3.60
	306702	AI022565	Hs.307670	EST	1.47	1.19
	306728	AI027359	Hs.272572	hemoglobin, alpha 2	1.28	2.83
	306751	AI032589		gb:ow70h12.s1 Soares_fetal_liver_spleen_	3.91	5.21
20	306767	AI038963	Hs.249118	ESTs	3.33	6.06
	306892	AI092465		gb:qa75h12.x1 Soares_fetal_heart_NbHH19W	3.77	7.46
	306897	AI093957		gb:qa33c06.s1 Soares_NhHMPu_S1 Homo sapi	2.12	2.85
	306956	AI125111		gb:am66f03.s1 Barstead spleen HPLRB2 Hom	6.10	10.52
	306958	AI125152		gb:am55e09.x1 Johnston frontal cortex Ho	1.72	1.56
25	307035	AI142774	Hs.119122	ribosomal protein L13a	2.00	4.70
	307041	AI144243		gb:qb95b12.x1 Soares_fetal_heart_NbHH19W	9.12	12.56
	307091	AI167439		gb:ox70h06.s1 Soares_NhHMPu_S1 Homo sapi	4.88	8.52
	307181	AI189251		gb:qc99g06.x1 Soares_pregnant_uterus_NbH	3.55	6.44
	307297	AI205798	Hs.111334	ferritin, light polypeptide	2.46	4.65
30	307317	AI208303	Hs.147333	EST	5.64	10.13
	307327	AI214142	Hs.246381	CD68 antigen	3.18	5.15
	307382	AI223158	Hs.147885	ESTs	2.02	3.73
	307410	AI241715	Hs.77039	ribosomal protein S3A	0.72	0.48
	307415	AI242118		gb:qh92b02.x1 Soares_NFL_T_GBC_S1 Homo s	2.38	3.51
35	307423	AI243206	Hs.179573	collagen, type I, alpha 2	2.60	5.44
	307426	AI243364		gb:qh30g11.x1 Soares_NFL_T_GBC_S1 Homo s	3.18	7.67
	307517	AI275055		gb:qj72d03.x1 Soares_NhHMPu_S1 Homo sapi	1.00	1.00
	307551	AI281556		gb:qu52f11.x1 NCL_CGAP_Lym6 Homo sapiens	3.40	11.20
	307561	AI282207		gb:qp65a12.x1 Soares_fetal_lung_NbHL19W	4.74	15.51
40	307608	AI290295		gb:qm01f02.x1 Soares_NhHMPu_S1 Homo sapi	3.50	7.19
	307657	AI306428	Hs.298262	ribosomal protein S19	1.76	2.44
	307691	AI318285		gb:tb17b01.x1 NCL_CGAP_Ov37 Homo sapiens	1.59	1.31
	307701	AI318583	Hs.276672	EST, Weakly similar to RL6_HUMAN 60S RI	1.90	2.13
	307718	AI333406	Hs.83753	small nuclear ribonucleoprotein polypept	0.45	0.99
45	307730	AI336092		gb:qx43b07.x1 Soares_fetal_lung_NbHL19W	1.51	0.99
	307760	AI342387		gb:qt27f07.x1 Soares_pregnant_uterus_NbH	1.00	1.00
	307764	AI342731		gb:qp26a07.x1 NCL_CGAP_Lu5 Homo sapiens	4.52	12.58
	307783	AI347274		gb:td05d02.x1 NCL_CGAP_Co16 Homo sapiens	1.42	1.00
	307796	AI350556		gb:qt18f09.x1 NCL_CGAP_GC4 Homo sapiens	6.57	9.61
50	307807	AI351799		gb:qt09d02.x1 NCL_CGAP_GC4 Homo sapiens	3.38	7.68
	307808	AI351826		gb:qt09g03.x1 NCL_CGAP_GC4 Homo sapiens	0.33	0.86
	307820	AI355761		gb:qt94a11.x1 NCL_CGAP_Co14 Homo sapiens	7.94	21.57
	307830	AI358722	Hs.276737	EST, Weakly similar to RSHU22 ribosomal	2.05	3.32
	307852	AI365541		gb:qz08g05.x1 NCL_CGAP_CLL1 Homo sapiens	3.18	5.21
55	307902	AI380462		gb:tg02h05.x1 NCL_CGAP_CLL1 Homo sapiens	3.13	4.99
	307997	AI434512	Hs.181165	eukaryotic translation elongation factor	1.00	3.01
	308002	AI435240	Hs.283442	ESTs	5.86	12.64
	308011	AI439473		gb:ti60a08.x1 NCL_CGAP_Lym12 Homo sapien	3.79	5.83
	308023	AI452732	Hs.251577	hemoglobin, alpha 1	0.38	0.88
60	308041	AI458824	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.36	6.06
	308059	AI468938	Hs.276877	EST, Weakly similar to RL10_HUMAN 60S R	1.80	1.98
	308085	AI474135	Hs.181165	eukaryotic translation elongation factor	3.38	4.14
	308101	AI475950	Hs.181165	eukaryotic translation elongation factor	1.30	3.87
	308106	AI476803		gb:ij77e12.x1 Soares_NSF_F8_9W_OT_PA_P_S2.38	8.72	8.72
65	308122	AI480123	Hs.309411	EST	2.70	3.86
	308154	AI500600		gb:tn93d08.x1 NCL_CGAP_UI2 Homo sapiens	0.66	1.33
	308171	AI523532	Hs.298766	ESTs, Weakly similar to schlafen4 [Mmu	2.48	4.86
	308211	AI557029	Hs.278572	anaplastic lymphoma kinase (K1-1)	2.43	2.14
	308213	AI557041		gb:PT2.1_12_E04.r tumor2 Homo sapiens cD	3.34	3.79
70	308216	AI557135		gb:PT2.1_13_H06.r tumor2 Homo sapiens cD	4.61	4.78
	308219	AI557246		gb:PT2.1_15_D07.r tumor2 Homo sapiens cD	4.87	7.94
	308271	AI567844	Hs.252259	ribosomal protein S3	2.40	6.35
	308319	AI583983	Hs.181165	eukaryotic translation elongation factor	2.45	3.33
	308362	AI613519	Hs.105749	KIAA0553 protein	1.24	1.41
75	308413	AI636253	Hs.196511	ESTs	3.16	4.82
	308450	AI650860	Hs.96840	KIAA1527 protein	1.79	2.68
	308464	AI672425	Hs.277117	EST, Moderately similar to I38055 myosi	4.87	8.27
	308588	AI718299		gb:as51g12.x1 Barstead aorta HPLRB6 Homo	3.90	5.64
	308599	AI719893		gb:as47d07.x1 Barstead aorta HPLRB6 Homo	3.32	5.12
80	308615	AI738593	Hs.101774	hypothetical protein FLJ23045	3.11	2.36
	308643	AI745040		gb:tr19a12.x1 NCL_CGAP_Ov23 Homo sapiens	3.98	3.69
	308673	AI760864		gb:wi09c10.x1 NCL_CGAP_CLL1 Homo sapiens	0.82	0.99
	308697	AI767143		gb:wi97a07.x1 NCL_CGAP_Kid12 Homo sapien	2.76	5.59
	308762	AI807405	Hs.259408	ESTs	3.17	6.30
85	308778	AI811109		gb:tr04c11.x1 NCL_CGAP_Ov23 Homo sapiens	1.00	1.00
	308782	AI811767	Hs.2186	eukaryotic translation elongation factor	2.94	5.15
	308808	AI818289		gb:wk52c01.x1 NCL_CGAP_Pr22 Homo sapiens	4.41	8.34
	308823	AI824118	Hs.217493	annexin A2	1.85	1.92
	308875	AI832332		gb:at48g03.x1 Barstead colon HPLRB7 Homo	2.52	3.80

	308879	AI832763	Hs.75958	thymosin, beta 4, X chromosome	3.38	7.96
	308886	AI833240		gb:at76d10.x1 Barstead colon HPLR87 Homo	3.06	2.65
	308898	AI858845		gb:w132d10.x1 NCI_CGAP_Ut1 Homo sapiens	2.45	3.44
5	308934	AI855023	Hs.177	phosphatidylinositol glycan, class H	4.14	6.76
	308966	AI870704		gb:w147h01.x1 NCI_CGAP_Ut1 Homo sapiens	1.00	1.00
	308979	AI873111		gb:w152h05.x1 NCI_CGAP_Bm25 Homo sapien	7.15	11.10
	309045	AI910902		gb:1q39i01.x1 NCI_CGAP_Ut1 Homo sapiens	0.61	0.59
	309051	AI911975		gb:w178d01.x1 NCI_CGAP_Lu24 Homo sapiens	1.78	4.42
10	309069	AI917366	Hs.78202	SWI/SNF related, matrix associated, act	3.27	5.88
	309083	AI922426	Hs.119598	ribosomal protein L3	2.39	3.34
	309105	AI925503	Hs.265884	ESTs	5.54	17.78
	309122	AI928178		gb:wo95a11.x1 NCI_CGAP_Kid11 Homo sapien	1.00	2.92
	309128	AI928816	Hs.180842	ribosomal protein L13	1.38	5.55
15	309164	AI937761		gb:wp84b09.x1 NCI_CGAP_Bm25 Homo saplen	2.43	3.11
	309177	AI951118		gb:wx63g05.x1 NCI_CGAP_Br18 Homo sapiens	0.81	0.97
	309288	AI991525	Hs.299426	ESTs	4.85	7.46
	309299	AW003478		gb:wq66c06.x1 NCI_CGAP_GC6 Homo sapiens	4.36	9.43
	309303	AW004823		gb:ws93a08.x1 NCI_CGAP_Co3 Homo sapiens	2.88	7.54
20	309411	AW085201	Hs.244144	EST	4.30	7.14
	309437	AW090702	Hs.278242	tubulin, alpha, ubiquitous	2.49	3.11
	309459	AW117645	Hs.65114	keratin 18	2.88	4.55
	309476	AW129368		gb:xe14b05.x1 NCI_CGAP_Ut4 Homo sapiens	2.08	6.60
	309499	AW136325	Hs.279771	Homo sapiens clone PP1596 unknown mRNA	2.82	3.55
	309529	AW150807	Hs.181357	laminin receptor 1 (67kD, ribosomal pro	4.78	3.95
25	309532	AW151119		gb:xg33e10.x1 NCI_CGAP_Ut1 Homo sapiens	1.18	4.40
	309526	AW192004	Hs.297681	serine (or cysteine) proteinase inhibit	4.46	12.06
	309541	AW194230	Hs.253100	EST, Moderately similar to GH1U Ig gamm	1.47	1.39
	309575	AW205581	Hs.253506	EST, Moderately similar to ATPN_HUMAN A	5.68	15.20
	309593	AW237221	Hs.181357	laminin receptor 1 (67kD, ribosomal prot	1.00	1.00
30	309595	AW238011	Hs.295605	mannosidase, alpha, class 2A, member 2	5.45	9.61
	309700	AW241170	Hs.179661	tubulin, beta polypeptide	1.41	1.25
	309747	AW264889		gb:xq36h02.x1 NCI_CGAP_Lu28 Homo sapiens	5.00	8.35
	309769	AW272346		gb:xs13c10.x1 NCI_CGAP_Kid11 Homo saplen	5.76	11.90
35	309782	AW275155	Hs.156110	immunoglobulin kappa constant	0.42	0.69
	309783	AW275401	Hs.254798	EST	1.00	4.11
	309799	AW276964		gb:xp58h01.x1 NCI_CGAP_Ov39 Homo sapiens	1.68	1.44
	309866	AW299916		gb:xs44c01.x1 NCI_CGAP_Kid11 Homo saplen	3.02	5.04
	309903	AW339071	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.05	1.18
40	309923	AW340684		gb:hd05g08.x1 Soares_NFL_T_GBC_S1 Homo s	2.30	3.67
	309928	AW341418		gb:hd08c03.x1 Soares_NFL_T_GBC_S1 Homo s	7.41	13.71
	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s	1.20	12.70
	309933	AW341936		gb:hb73f10.x1 NCI_CGAP_Ut2 Homo sapiens	4.90	18.29
	309964	AW449111	Hs.257111	hypothetical protein MGC3265	1.99	3.07
45	310002	AI439096	Hs.323079	Homo sapiens mRNA; cDNA DKFZp564P116 (fr	0.20	0.47
	310096	AW136822	Hs.172824	ESTs, Weakly similar to B48013 proline-r	1.51	1.22
	310098	AI685841	Hs.161354	ESTs	0.31	0.76
	310109	AI203094	Hs.148633	ESTs	2.06	5.83
	310112	AW197233	Hs.147253	ESTs	2.92	3.55
	310115	AI811317	Hs.223796	ESTs	1.25	0.84
50	310121	AW195642	Hs.148901	ESTs	1.00	2.71
	310146	AI206614	Hs.197422	ESTs	9.50	15.31
	310193	AI627653	Hs.147562	ESTs	2.85	4.18
	310255	AW450439	Hs.153378	ESTs	4.26	10.63
55	310261	AI240483	Hs.201217	ESTs	3.28	4.40
	310264	AI915771	Hs.74170	metallothionein 1E (functional)	0.26	0.86
	310275	AI242102	Hs.213636	ESTs	5.43	8.19
	310282	AI243332	Hs.156055	ESTs	3.15	8.06
	310290	AW013815	Hs.149103	ESTs	2.19	3.12
60	310333	AI253200	Hs.145402	ESTs	1.17	1.91
	310346	AI261340	Hs.145517	ESTs	4.81	9.95
	310385	AI263392	Hs.156151	ESTs	5.96	7.79
	310443	AW119018	Hs.164231	ESTs	2.90	4.63
	310444	AW196632	Hs.252956	ESTs	0.85	1.01
65	310446	AI275715	Hs.145926	ESTs	2.18	3.85
	310468	AI984074	Hs.196398	ESTs	3.39	5.19
	310477	AI948801	Hs.171073	ESTs	1.00	1.00
	310512	AW275603	Hs.200712	ESTs	3.87	8.12
	310514	AI681145	Hs.160724	ESTs	3.30	7.33
70	310524	AW082270	Hs.12496	ESTs, Highly similar to AC004836 1 simil	0.72	1.44
	310547	AI302654	Hs.208024	ESTs	3.26	3.46
	310584	AI653007	Hs.156304	ESTs	2.39	4.08
	310608	AI962234	Hs.196102	ESTs	5.60	6.49
	310624	AI341594		gb:Human endogenous retrovirus H proteas	4.91	9.09
75	310636	AI814373	Hs.164175	ESTs	1.85	1.71
	310648	AI347863	Hs.156672	ESTs	0.17	0.69
	310694	AI654370	Hs.157752	Homo sapiens mRNA full length insert cDN	5.40	13.22
	310695	AI472124	Hs.157757	ESTs	4.82	6.27
	310714	AI418446	Hs.157882	ESTs	1.76	3.51
	310722	AI989803	Hs.157289	ESTs	1.14	6.85
80	310756	AI916560	Hs.158707	ESTs	8.46	13.01
	310764	AI376769	Hs.167172	ESTs	4.76	7.37
	310848	AI459554	Hs.161286	ESTs	2.84	1.96
	310851	AW291714	Hs.221703	ESTs	1.00	2.32
85	310854	AI421677	Hs.161332	ESTs	6.37	7.94
	310858	AI871000	Hs.161330	ESTs	6.07	9.84

5	310864	AI924558	Hs.161399	ESTs	0.87	0.78
	310875	T47764	Hs.132917	ESTs	1.00	3.63
	310896	AW157731	Hs.270982	ESTs, Moderately similar to ALU7_HUMAN A	7.07	16.68
	310922	AW195634	Hs.170401	ESTs	1.00	1.00
	310955	AI560210	Hs.263912	ESTs	10.08	17.66
10	310957	AW190974	Hs.196918	ESTs	2.18	3.18
	311000	AI521830	Hs.171050	ESTs	3.08	6.64
	311012	AW288070	Hs.241097	ESTs	1.23	3.77
	311034	AI564023	Hs.311389	ESTs, Moderately similar to PT0375 natur	2.44	2.09
	311074	AW290922	Hs.199848	ESTs	6.04	14.19
15	311134	AI990849	Hs.196971	ESTs	3.54	6.96
	311174	AW450552	Hs.205457	perixin	0.65	0.95
	311187	AI638374	Hs.224189	ESTs	2.46	2.78
	311220	AI656040	Hs.196532	ESTs	1.10	2.52
	311230	AI989808	Hs.197663	ESTs	1.41	1.75
20	311236	AI653378	Hs.197674	ESTs	2.18	2.11
	311242	AW016812	Hs.200266	ESTs	0.63	5.11
	311258	AI671221	Hs.198887	ESTs	1.00	1.41
	311277	AW072813	Hs.270868	ESTs, Moderately similar to ALU4_HUMAN A	2.56	1.94
	311294	AA826425	Hs.291829	ESTs	1.04	2.69
25	311308	F12664	Hs.49000	ESTs	1.96	6.70
	311351	AI682303	Hs.201274	ESTs	4.77	9.38
	311390	AW392997	Hs.202280	ESTs	2.80	6.06
	311405	AW290961	Hs.201815	ESTs	3.80	11.66
	311409	AI698839		gb:wd31f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.84	6.94
30	311420	AI936291	Hs.209867	ESTs	5.30	12.56
	311443	AI791521	Hs.192206	ESTs	4.39	6.09
	311467	AI934909	Hs.175377	ESTs	1.00	1.04
	311479	AI933672	Hs.211399	ESTs	2.76	5.61
	311488	R57390	Hs.301064	araplin 1	2.50	5.73
35	311495	AW300077	Hs.221358	ESTs	3.63	6.09
	311511	AW444568	Hs.210303	ESTs	2.00	2.87
	311534	AW130351	Hs.243549	ESTs	0.31	1.33
	311537	AI805121	Hs.211828	ESTs	3.69	5.85
	311543	AI681360	Hs.201269	ESTs	1.73	1.34
40	311551	AW449774	Hs.296380	POM (POM121 rat homolog) and ZP3 fusion	3.31	6.12
	311557	AI819230	Hs.211238	Interleukin-1 homolog 1	1.00	1.00
	311558	Z44432	Hs.63128	KIAA1292 protein	2.25	3.41
	311559	AW008271	Hs.265848	similar to rat myomegalin	2.68	5.90
	311563	AI922143	Hs.211334	ESTs	2.39	3.32
45	311586	AI827834	Hs.211227	ESTs	2.47	3.85
	311616	AW450675	Hs.212709	ESTs	1.00	1.00
	311621	AI924307	Hs.213464	ESTs	4.16	6.74
	311635	AI928456	Hs.213081	ESTs	2.17	3.76
	311668	AW193674	Hs.240044	ESTs	2.60	3.12
50	311672	R11807	Hs.20914	hypothetical protein FLJ23056	2.79	5.18
	311683	AW183738	Hs.232644	ESTs	0.19	0.96
	311700	R49601	Hs.171495	retinoic acid receptor, beta	6.28	8.83
	311714	AW131785	Hs.246831	ESTs, Weakly similar to CLK_HUMAN VOLTA	5.00	8.17
	311735	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	0.96	0.72
55	311743	T99079	Hs.191194	ESTs	1.00	1.95
	311783	AI682478	Hs.13528	hypothetical protein FLJ14054	0.16	0.77
	311785	AI056769	Hs.133512	ESTs	1.34	3.97
	311799	AA780791	Hs.14014	ESTs, Weakly similar to KIAA0973 protein	8.52	13.32
	311819	AW265275	Hs.254325	ESTs	3.58	3.91
60	311823	AI089422	Hs.131297	ESTs	1.40	1.72
	311877	AA349893	Hs.85339	G protein-coupled receptor 39	0.95	0.91
	311886	AA522738	Hs.132554	ESTs	0.88	0.87
	311896	AW206447		gb:U1-H-B11-atg-g-02-0-U1.s1 NCL_CGAP_Su	1.66	1.13
	311910	N28365	Hs.22579	Homo sapiens clone CDABP0036 mRNA sequen	1.66	2.30
65	311923	T60843	Hs.189679	ESTs	0.42	2.63
	311933	AI597963	Hs.118726	ESTs	1.88	3.02
	311959	T67282	Hs.124733	ESTs	2.02	2.33
	311960	AW440133	Hs.189690	ESTs	3.87	6.62
	311967	AI382726	Hs.182434	ESTs	5.80	8.14
70	311975	AA804374	Hs.272203	Homo sapiens cDNA FLJ20843 fis, clone AD	0.98	3.26
	312005	T78450	Hs.13941	ESTs	0.12	1.39
	312028	T78886	Hs.284450	ESTs	3.78	4.92
	312046	AI580018	Hs.268591	ESTs	4.11	7.32
	312056	T83748	Hs.268594	ESTs	2.36	3.08
75	312064	AA676713	Hs.191155	ESTs	3.34	5.28
	312088	AW303760	Hs.13685	ESTs	1.60	1.15
	312093	T91809	Hs.121296	ESTs	0.68	0.85
	312094	Z78390		gb:HSZ78390 Human fetal brain S. Meier-E	3.05	4.48
	312097	AI352096	Hs.112180	zinc finger protein 148 (pH2-52)	4.52	9.70
80	312118	T85332	Hs.178294	ESTs	2.40	2.60
	312128	AI052609	Hs.17631	Homo sapiens cDNA FLJ20118 fis, clone CO	2.39	3.53
	312147	T89855	Hs.195648	ESTs	0.67	1.03
	312175	AA953383	Hs.127554	ESTs	5.85	10.60
	312179	AI052572	Hs.269864	ESTs	2.41	3.32
85	312201	AI926365	Hs.91139	solute carrier family 1 (neuronal/epitha	0.24	0.89
	312207	H90213	Hs.191330	ESTs	2.20	4.55
	312220	N74613		gb:z55a07.s1 Soares fetal liver spleen	4.28	11.13
	312252	AI128388	Hs.143655	ESTs	1.64	1.57
	312304	AA491849	Hs.269392	ESTs	0.12	2.47

	312318	AW235092	Hs.143981	ESTs	3.46	5.69
	312319	AA216598	Hs.180780	TERA protein	5.78	4.46
	312321	R65210	Hs.186937	ESTs	0.44	1.74
5	312331	AA825512	Hs.289101	glucose regulated protein, 58kD	3.73	5.96
	312339	AA524394	Hs.165544	ESTs	3.07	0.95
	312353	AI675558	Hs.181867	ESTs	10.08	16.73
	312375	AI375096	Hs.172405	cell division cycle 27	2.78	3.71
	312376	R52089	Hs.172717	ESTs	1.00	1.00
10	312389	AI863140		gb:tz43h12.x1 NCI_CGAP_Bm52 Homo sapien	2.37	3.98
	312437	AA995028		gb:RC4-BT0629-120200-011-b10 BT0629 Homo	4.06	5.41
	312440	AI051133	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	1.00	1.00
	312451	R59989	Hs.176539	ESTs	4.96	10.04
	312458	AI167637	Hs.146924	ESTs	1.11	1.00
15	312507	AI168177	Hs.143653	ESTs	5.89	8.24
	312520	AI742591	Hs.205392	ESTs	3.30	8.92
	312548	AI566228	Hs.159426	hypothetical protein PRO2121	1.38	1.65
	312564	H21520	Hs.35088	ESTs	0.40	0.77
	312583	AI193122	Hs.124141	ESTs	0.13	0.94
20	312599	AI865073	Hs.125720	ESTs	3.75	5.29
	312602	AA046451	Hs.165200	ESTs	6.78	12.93
	312645	H52121	Hs.193007	ESTs	0.38	1.13
	312666	AI240582	Hs.214678	ESTs	0.98	2.03
	312689	AW450461	Hs.203965	ESTs	0.21	0.61
	312817	H75459	Hs.233425	ESTs	1.51	0.85
25	312846	AW152104	Hs.200879	ESTs	8.93	13.78
	312873	AI690071	Hs.283552	ESTs, Weakly similar to unnamed protein	4.20	6.23
	312893	AI016204	Hs.172922	ESTs	2.67	3.15
	312902	AW292797	Hs.130316	ESTs, Weakly similar to T2D3_HUMAN TRANS	1.19	0.71
30	312925	N90868	Hs.271695	ESTs	2.60	4.25
	312936	AI681581	Hs.121525	ESTs	1.00	1.17
	312975	AI640506	Hs.293119	ESTs, Weakly similar to ALU7_HUMAN ALU S	2.30	4.80
	312978	N24887	Hs.292500	ESTs	0.80	1.05
	312980	AA497043	Hs.115685	ESTs	3.12	3.60
35	312984	N25871	Hs.177337	ESTs	2.03	2.13
	313000	AI147412	Hs.146657	ESTs	5.52	8.42
	313029	AA731520	Hs.170504	ESTs	0.96	1.39
	313039	AI419290	Hs.149990	ESTs, Weakly similar to unnamed protein	6.48	13.20
40	313049	AW293055	Hs.119357	ESTs	6.44	10.73
	313056	AI651930	Hs.135684	ESTs	1.51	2.04
	313058	D81015	Hs.125382	ESTs	0.25	1.50
	313070	AI422023	Hs.161338	ESTs	8.56	11.60
	313087	AI676164	Hs.204339	ESTs	3.72	4.56
	313130	AW449171	Hs.168677	ESTs	3.28	5.06
45	313136	N59284	Hs.288010	ESTs	0.49	1.36
	313153	AI240838	Hs.132750	ESTs	5.36	5.52
	313210	N74077	Hs.197043	ESTs	0.30	0.66
	313236	AW236169	Hs.83513	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.16	8.76
	313239	W19632	Hs.124170	ESTs	1.00	3.87
50	313265	N93466	Hs.121764	ESTs, Weakly similar to testicular tekf	0.74	2.06
	313267	AI770008	Hs.129583	ESTs	0.23	1.30
	313275	AI027604	Hs.159650	ESTs	6.68	9.57
	313290	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	1.34	1.07
	313292	AI362991	Hs.202121	ESTs, Weakly similar to env protein [H.s	2.00	4.32
55	313325	AI420511	Hs.127832	ESTs	1.20	2.27
	313357	AW074848	Hs.201501	ESTs	4.02	5.33
	313393	AI674685	Hs.200141	ESTs	1.36	2.84
	313399	AW376889	Hs.194097	ESTs	2.58	5.26
	313414	AI241540	Hs.132933	ESTs	6.57	15.07
60	313417	AA741151	Hs.137323	ESTs	0.63	3.01
	313457	AA576052	Hs.193223	Homo sapiens cDNA FLJ11646 fis, clone HE	2.78	4.70
	313499	AI261390	Hs.146085	KIAA1345 protein	0.91	2.37
	313516	AA029058	Hs.135145	ESTs	3.41	7.08
	313556	AA628517	Hs.118502	ESTs	0.23	0.70
65	313569	AI273419	Hs.135146	hypothetical protein FLJ13984	1.88	1.00
	313570	AA041455	Hs.209312	ESTs	0.73	2.27
	313638	AI753075	Hs.104627	Homo sapiens cDNA FLJ10158 fis, clone HE	1.00	1.72
	313662	AA740151	Hs.130425	ESTs	0.20	1.42
	313671	W49823	Hs.104613	RP42 homolog	1.00	1.00
70	313672	AW468891	Hs.122948	ESTs	3.46	5.80
	313690	AI493591	Hs.78146	platelet/endothelial cell adhesion molec	0.51	0.97
	313711	AA398070	Hs.133471	ESTs	0.18	1.01
	313723	AA070412		gb:zm68c10.s1 Stratagene neuroepithelium	1.08	1.03
	313726	AI744687	Hs.257806	ESTs	2.13	2.99
75	313774	AW136836	Hs.144583	ESTs	1.38	1.19
	313784	AA910514	Hs.134905	ESTs	3.88	5.78
	313790	AW078569	Hs.177043	ESTs	0.22	2.06
	313832	AW271022	Hs.133294	ESTs	1.15	0.91
	313834	AW418779	Hs.114889	ESTs	0.68	3.14
80	313835	AI538438	Hs.159087	ESTs	5.74	8.88
	313852	H18633	Hs.123641	protein tyrosine phosphatase, receptor t	0.16	1.14
	313854	AW470806	Hs.275002	ESTs	2.09	4.05
	313865	AA731470	Hs.163839	ESTs	3.41	4.09
	313871	AW471088	Hs.145950	ESTs	5.28	6.83
85	313883	AI949384		gb:nu76d01.s1 NCI_CGAP_Alv1 Homo sapiens	2.90	10.91
	313915	AI969390	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	1.00	1.00

	313926	AW473830	Hs.171442	ESTs	3.40	4.11
	313948	AW452823	Hs.135268	ESTs	5.77	9.15
	313978	AI870175	Hs.13957	ESTs	0.46	0.75
5	313983	AI829133	Hs.226780	ESTs	4.10	6.40
	314035	AA164199	Hs.270152	ESTs	5.88	7.90
	314037	AW300048	Hs.275272	ESTs	1.00	3.79
	314040	AA166970	Hs.118748	ESTs	7.60	11.33
	314067	AW293538	Hs.51743	KIAA1340 protein	1.86	1.21
10	314103	AI028477	Hs.132775	ESTs	2.90	5.29
	314107	AA806113	Hs.189025	ESTs	2.00	1.66
	314113	AA218986	Hs.118854	ESTs	0.91	4.17
	314124	AW118745	Hs.9460	Homo sapiens mRNA; cDNA DKFZp547C244 (fr	2.53	3.32
	314126	AA226431	gbnc18b12.s1 NCI_CGAP_Pr1 Homo sapiens	3.13	5.08	
15	314128	AA935633	Hs.194628	ESTs	2.90	6.35
	314151	AA236163	Hs.202430	ESTs	4.15	6.45
	314184	AW081795	Hs.233465	ESTs	3.44	4.65
	314192	AW290975	Hs.118923	ESTs	1.00	1.23
	314244	AL036450	Hs.103238	ESTs	2.88	3.67
20	314253	AA278679	Hs.189510	ESTs	4.98	7.16
	314262	AW086215	Hs.246096	ESTs	0.38	1.94
	314320	AA811598	Hs.275809	ESTs	3.34	5.66
	314332	AL037551	Hs.95612	ESTs	2.85	2.09
	314335	AA287443	Hs.142570	Homo sapiens clone 24629 mRNA sequence	4.35	4.78
	314340	AW304350	Hs.130879	ESTs, Moderately similar to putative p15	0.77	0.86
25	314351	AA292275	Hs.193746	ESTs	3.07	3.77
	314376	AI628633	Hs.324679	ESTs	4.10	6.11
	314443	AA827125	Hs.192043	ESTs	6.20	13.67
	314458	AI217440	Hs.143873	ESTs	0.58	2.49
	314466	AA767818	Hs.122707	ESTs	2.53	2.62
30	314478	AI521173	Hs.125507	DEAD-box protein	3.94	5.65
	314482	AL043807	Hs.134182	ESTs	1.30	1.44
	314506	AA833655	Hs.206868	Homo sapiens cDNA FLJ14056 fis, clone HE	3.28	3.47
	314519	R42554	Hs.210862	T-box, brain, 1	3.12	6.16
35	314529	AL046412	Hs.202151	ESTs	3.43	6.87
	314546	AW007211	Hs.16131	hypothetical protein FLJ12876	1.38	1.00
	314562	AI564127	Hs.143493	ESTs	2.29	5.27
	314579	AW197442	Hs.116998	ESTs	3.87	5.75
	314580	AW451832	Hs.255938	ESTs, Moderately similar to KIAA1200 pro	0.10	0.71
40	314585	AA918474	Hs.216363	ESTs	1.08	1.40
	314589	AW384790	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	1.00	1.00
	314592	AA435761	Hs.192148	ESTs	0.90	2.60
	314603	AA418024	Hs.270670	ESTs	4.56	6.29
	314604	AA946582	Hs.8700	deleted in liver cancer 1	3.42	3.92
45	314606	AA418241	Hs.188767	ESTs	2.97	4.55
	314648	AA878419		gb:EST391378 MAGE resequences, MAGP Homol.42		1.36
	314699	AI038719	Hs.132801	ESTs	3.66	4.97
	314701	AI754634	Hs.131987	ESTs	0.03	0.90
	314710	AI669131	Hs.290989	EST	3.40	7.52
50	314750	AI095005	Hs.135174	ESTs	2.80	6.54
	314767	AW135412	Hs.164002	ESTs	3.20	4.26
	314801	AA481027	Hs.109045	hypothetical protein FLJ10498	1.00	1.00
	314817	AI694139	Hs.192855	ESTs	0.91	0.99
	314835	AI281370	Hs.76064	ribosomal protein L27a	5.75	7.44
55	314852	AI903735		gb:MR-BT035-200199-031 BT035 Homo sapien	1.68	4.34
	314853	AA729232	Hs.153279	ESTs	0.60	1.85
	314940	AW452768	Hs.162045	ESTs	10.10	16.20
	314941	AA515902	Hs.130650	ESTs	0.31	1.02
	314943	AI476797	Hs.184572	cell division cycle 2, G1 to S and G2 to	2.18	0.37
60	314955	AA521382	Hs.192534	ESTs	2.59	3.90
	314973	AW273128	Hs.300268	ESTs	1.05	1.25
	315004	AA527941	Hs.325351	EST	5.64	13.63
	315006	AI538613	Hs.298241	Transmembrane protease, serine 3	0.52	1.78
	315033	AI493046	Hs.146133	ESTs	2.46	1.00
65	315035	AI569476	Hs.177135	ESTs	0.34	1.33
	315056	AI202703	Hs.152414	ESTs	2.10	2.64
	315069	AI821517	Hs.105866	ESTs	1.00	1.30
	315071	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 fis, clone C	1.78	1.00
	315073	AW452948	Hs.257631	ESTs	1.17	1.52
70	315078	AA568548	Hs.190616	ESTs	3.00	3.79
	315080	AA744550	Hs.136345	ESTs	1.00	1.00
	315120	AA564991	Hs.269477	ESTs	0.64	1.44
	315175	AI025842	Hs.152530	ESTs	0.61	1.91
	315193	AI241331	Hs.131765	ESTs	1.06	0.97
75	315196	AA972756	Hs.44898	Homo sapiens clone TCCTA00151 mRNA sequ	0.48	1.96
	315200	AI808235	Hs.307686	EST	3.76	9.40
	315254	AI474433	Hs.179556	ESTs	5.37	9.36
	315353	AW452608	Hs.279610	hypothetical protein FLJ10493	1.00	1.30
	315397	AA218940	Hs.137516	fidgulin-like 1	3.38	2.24
80	315403	AW362980	Hs.163924	ESTs	2.04	5.23
	315431	AA622104	Hs.184838	ESTs	2.36	8.04
	315454	AI239473		gb:qh36f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.46	7.64
	315455	AW393391	Hs.156919	ESTs	3.78	5.76
	315473	AI681671	Hs.312671	ESTs, Moderately similar to OVCA1	0.89	2.15
85	315483	AW512763	Hs.222024	transcription factor BMAL2	2.32	1.96
	315526	AI193048	Hs.128685	ESTs	1.67	1.78

	315530	AI200852	Hs.127780	ESTs	1.05	1.01
	315541	AI168233	Hs.123159	sperm associated antigen 4	0.85	0.56
	315552	AW445034	Hs.256578	ESTs	1.00	2.22
5	315562	AA737415	Hs.152826	ESTs	2.66	2.48
	315577	AW513545	Hs.17283	hypothetical protein FLJ10890	2.20	2.25
	315587	AI268399	Hs.140489	ESTs	1.00	1.04
	315589	AW072387	Hs.158258	Homo sapiens mRNA; cDNA DKFZp434B1272 (f	0.14	1.05
	315623	AA364078	Hs.258189	ESTs	7.44	12.56
10	315634	AA837085	Hs.220585	ESTs	0.50	1.40
	315668	AA912347	Hs.136585	ESTs	0.43	1.22
	315677	AI932662	Hs.164073	ESTs	0.60	1.39
	315708	AW440742	Hs.155556	hypothetical protein FLJ20202	2.18	3.77
	315707	AI418055	Hs.161160	ESTs	2.88	2.63
15	315730	H25899	Hs.201591	ESTs	0.11	0.60
	315745	AI821759	Hs.191856	ESTs	3.50	7.25
	315791	AA678177	gb:z15a05.s1 Soares_fetal_liver_spleen_	1.78	2.63	
	315801	AA827752	Hs.286134	ESTs	4.31	6.23
	315820	AI652022	Hs.258785	ESTs	2.35	3.01
20	315878	AA583336	Hs.189046	ESTs	2.12	2.64
	315905	AI821911	Hs.209452	ESTs	1.03	1.97
	315923	AI052789	Hs.133263	ESTs	2.63	5.06
	315954	AW276810	Hs.254859	ESTs, Moderately similar to ALU5_HUMAN A	1.21	0.85
	315978	AA830893	Hs.119769	ESTs	3.09	3.41
25	316001	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	2.20	6.82
	316011	AW516953	Hs.201372	ESTs	0.35	1.63
	316012	AA764950	Hs.119898	ESTs	6.56	8.13
	316040	AI983409	Hs.189226	ESTs	5.69	10.69
	316048	AI720759	Hs.224971	ESTs	2.84	10.45
	316076	AW297895	Hs.116424	ESTs	0.30	1.05
30	316124	AI308862	Hs.167028	ESTs	1.00	1.43
	316151	AI806016	Hs.156520	ESTs	5.80	9.03
	316187	AW518299	Hs.192253	ESTs	1.20	3.96
	316204	AA731509	Hs.120257	ESTs	4.92	6.94
	316232	AW297853	Hs.251203	ESTs	1.48	1.60
35	316275	AI671041	Hs.292611	ESTs, Moderately similar to ALU1_HUMAN A	5.85	12.14
	316291	AW375974	Hs.156704	ESTs	2.73	2.69
	316303	AA740994	Hs.209609	ESTs	1.53	1.26
	316344	AA744518	Hs.120610	ESTs	3.66	8.34
	316346	AI028478	Hs.157447	ESTs	3.51	6.69
40	316365	AI627845	Hs.210776	ESTs	2.50	4.33
	316380	AI393378	Hs.164496	ESTs	1.16	2.16
	316470	AA809902	Hs.243813	ESTs	5.40	10.34
	316509	AA767310	Hs.291766	ESTs	2.48	2.89
	316514	AA768037	Hs.291671	ESTs	4.70	6.04
45	316519	AI929097	gb:cd10c11.s1 NCL_CGAP_GCB1 Homo sapiens	4.41	9.70	
	316609	AW292520	Hs.122082	ESTs	1.00	2.89
	316633	AI125586	Hs.127955	ESTs	2.61	3.72
	316700	AW172316	Hs.252961	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.46	4.64
50	316711	AI743721	Hs.285316	ESTs, Moderately similar to ALU7_HUMAN A	4.45	6.95
	316713	AI090671	Hs.134807	hypothetical protein FLJ12057	0.30	2.40
	316715	AA440266	Hs.170673	ESTs, Weakly similar to AF126780 1 refn	0.20	1.45
	316787	AW369770	Hs.130351	ESTs	4.05	5.53
	316809	AA825839	Hs.202238	ESTs	2.26	3.82
55	316811	AA922060	Hs.132471	ESTs	1.00	1.32
	316812	AW135045	Hs.232001	ESTs	3.28	4.70
	316818	AA827176	Hs.124316	ESTs	0.67	1.81
	316824	AA837416	Hs.124299	ESTs	3.53	6.00
	316827	AI380429	Hs.172445	ESTs	0.72	1.56
60	316891	AW296119	Hs.202536	ESTs	1.64	2.97
	316951	AA134365	Hs.57548	ESTs	1.45	1.08
	316970	AA860172	Hs.132406	ESTs	1.00	1.53
	316971	AA860212	Hs.170991	ESTs	1.08	1.96
	316990	AA861611	Hs.130643	ESTs	5.44	10.04
65	317001	AI627917	Hs.233694	hypothetical protein FLJ11350	3.56	4.37
	317008	AW051597	Hs.143707	ESTs	0.69	1.37
	317051	AA873253	Hs.126233	ESTs	6.18	12.72
	317128	AA971374	Hs.125674	ESTs	1.87	2.66
	317129	H12523	Hs.78521	Homo sapiens cDNA: FLJ21193 fis, clone C	4.12	6.64
70	317137	AW341567	Hs.125710	ESTs	2.82	5.12
	317196	AI348258	Hs.153412	ESTs	1.98	2.51
	317212	AI866468	Hs.148294	ESTs	1.86	2.83
	317223	AW297920	Hs.130054	ESTs	0.83	1.57
	317224	D56760	Hs.93029	sparc/osteoneclin, cwcv and kazal-like d	2.74	0.86
75	317266	AA906289	Hs.203614	ESTs	1.00	1.00
	317282	AI807444	Hs.176101	ESTs	2.60	4.21
	317285	AW370882	Hs.222080	ESTs	1.96	3.49
	317302	AA908709	Hs.135564	ESTs	7.16	8.32
	317304	AW449899	Hs.130184	ESTs	1.38	2.28
	317320	AA927151	Hs.130452	ESTs	3.58	8.13
80	317413	AW341701	Hs.126822	ESTs	2.08	4.92
	317417	AA918420	Hs.145378	ESTs	3.06	4.79
	317452	AA972965	Hs.135568	ESTs	4.22	9.21
	317519	AI859695	Hs.126860	ESTs	1.88	4.15
85	317521	AI824338	Hs.126891	ESTs	3.12	4.55
	317529	AI916517	Hs.126865	ESTs	2.73	3.34

5	317570	AI733361	Hs.127122	ESTs	1.00	2.43
	317571	AA938663	Hs.199828	ESTs	5.20	11.95
	317598	AW206035	Hs.192123	ESTs	0.33	1.56
	317627	AI346110	Hs.132553	ESTs	1.50	1.39
	317650	AI733310	Hs.127346	ESTs	0.48	1.46
10	317659	AA961216	Hs.127785	ESTs	4.18	7.14
	317674	AW294909	Hs.132208	ESTs	2.92	3.20
	317686	AA969051	Hs.187319	ESTs	1.00	1.01
	317692	AI307659	Hs.174794	ESTs	5.33	9.59
	317701	AI674774	Hs.128014	ESTs	1.00	1.00
15	317711	AI733015	Hs.272189	ESTs	5.13	7.81
	317722	AI733373	Hs.128119	ESTs	2.50	6.03
	317756	AA973687	Hs.128320	ESTs	1.59	1.30
	317777	AI143525	Hs.47313	KIAA0258 gene product	1.00	2.48
	317799	AI498273	Hs.128808	ESTs	1.78	2.11
20	317803	AA983251	Hs.128899	ESTs	0.80	1.08
	317821	AI368158	Hs.70983	PTPL1-associated RhoGAP 1	0.17	0.68
	317848	AI820575	Hs.129086	Homo sapiens cDNA FLJ12007 fis, clone HE	5.30	8.16
	317850	N29974	Hs.152982	hypothetical protein FLJ13117	1.30	2.28
	317861	AW341064	Hs.129119	ESTs	2.18	5.93
25	317865	AI298794	Hs.129130	ESTs	4.48	8.20
	317869	AW295184	Hs.129142	deoxyribonuclease II beta	0.44	0.99
	317881	AI827248	Hs.224398	Homo sapiens cDNA FLJ11469 fis, clone HE	4.06	2.23
	317890	AI915599	Hs.129225	ESTs	4.68	7.48
	317899	AI952430	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	3.14	3.37
30	317986	AI005163	Hs.201378	ESTs, Weakly similar to T12545 hypotheti	0.28	1.66
	318001	AW235697	Hs.130980	ESTs	5.12	9.97
	318016	AI016694	Hs.256921	ESTs	1.86	4.50
	318023	AW243058	Hs.131155	ESTs	2.92	5.22
	318054	AW449270	Hs.232140	ESTs	3.92	6.37
35	318068	AI024540	Hs.131574	ESTs	1.21	1.27
	318117	AI208304	Hs.250114	ESTs	0.86	1.17
	318187	AI792585	Hs.133272	ESTs, Weakly similar to ALUC_HUMAN IIII.	5.90	6.98
	318223	AI077540	Hs.134090	ESTs	1.05	0.90
	318240	AI085377	Hs.143610	ESTs	3.10	2.40
40	318255	AI082692	Hs.134662	ESTs	0.02	1.05
	318266	AI554341	Hs.271443	ESTs	6.12	10.55
	318330	AI093840	Hs.143758	ESTs	4.98	7.90
	318369	AI493501	Hs.170974	ESTs	2.48	5.62
	318428	AI949409	Hs.194591	ESTs	0.77	0.45
45	318458	AI149783	Hs.158438	ESTs	3.54	4.92
	318467	AI151395	Hs.144834	ESTs	4.56	5.62
	318473	AI939339	Hs.146883	ESTs	2.08	4.05
	318476	AI693927	Hs.265165	ESTs	4.22	8.07
	318487	AI167877	Hs.143716	ESTs	1.47	1.05
50	318488	AI217431	Hs.144709	ESTs	1.40	4.14
	318491	T26477	Hs.22883	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.84	1.90
	318499	T25451		gb:PTHI188 HTCDL1 Homo sapiens cDNA 5/3	2.58	5.20
	318537	AA377908	Hs.13254	ESTs	3.26	4.18
	318538	N28625	Hs.74034	Homo sapiens clone 24651 mRNA sequence	0.35	1.07
55	318547	R20578	Hs.90431	ESTs	3.22	4.60
	318552	R18364	Hs.90363	ESTs	4.87	9.06
	318575	R55102	Hs.107761	ESTs, Weakly similar to unnamed protein	1.91	1.98
	318580	T34571	Hs.49007	poly(A) polymerase alpha	2.74	6.22
	318587	AA779704	Hs.168830	Homo sapiens cDNA FLJ12136 fis, clone MA	0.85	2.45
60	318596	AI470235	Hs.172698	EST	4.88	4.93
	318622	T48325	Hs.237658	apolipoprotein A-II	4.80	12.51
	318629	N25163	Hs.8861	ESTs	0.39	1.04
	318637	AA243539	Hs.9196	hypothetical protein	1.72	3.57
	318648	T77141	Hs.184411	albumin	6.27	8.91
65	318650	AA393302	Hs.176626	hypothetical protein EDAG-1	3.96	8.84
	318671	AA188823	Hs.299254	Homo sapiens cDNA: FLJ23597 fis, clone L	1.53	0.81
	318679	T68115	Hs.10336	ESTs	1.00	2.19
	318711	AI936475	Hs.101282	Homo sapiens cDNA: FLJ21238 fis, clone C	3.05	3.18
	318725	AI962487	Hs.242890	ESTs	1.08	2.46
70	318728	Z30201	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S	0.77	1.33
	318740	NM_002543	Hs.77729	oxidized low density lipoprotein (lectin	0.25	1.49
	318776	R24963	Hs.23766	ESTs	1.00	3.01
	318784	H00148	Hs.5181	proliferation-associated 2G4, 38kD	2.70	3.86
	318816	F07873	Hs.21273	ESTs	3.90	7.13
75	318865	H10818		gb:ym04f10.r1 Soares infant brain 1N1B H	2.25	3.56
	318879	R55332	Hs.18268	adenylate kinase 5	1.78	5.00
	318881	Z43224	Hs.124952	ESTs	4.79	14.13
	318894	F08138	Hs.7387	DKFZP564B116 protein	5.31	7.00
	318901	AW368520	Hs.301528	L-tyrosine/alpha-amino acid amine	1.03	0.91
80	318925	Z43577	Hs.21470	ESTs	2.23	3.80
	318936	AI219221	Hs.308298	ESTs	1.86	7.16
	318982	Z44140	Hs.269622	ESTs	5.84	9.79
	318986	Z44186	Hs.169161	ESTs, Highly similar to MAON_HUMAN NADP-	1.00	1.00
	319041	Z44720	Hs.98365	ESTs, Weakly similar to weak similarity	3.38	6.11
85	319103	H05896	Hs.4993	KIAA1313 protein	1.00	1.07
	319170	R13678	Hs.285306	putative selenocysteine lyase	3.79	5.03
	319196	F07953	Hs.16085	putative G-protein coupled receptor	1.00	2.98
	319199	F07361	Hs.13308	ESTs	3.53	5.66
	319242	F11472	Hs.12639	ESTs	5.87	7.26

	319263	T65331	Hs.81360	Homo sapiens cDNA: FLJ21927 fis, clone H	1.81	1.57
	319267	F11802	Hs.6818	ESTs	1.10	4.72
	319270	R13474	Hs.290263	ESTs	4.80	10.40
5	319279	T65094	Hs.12677	CGI-147 protein	1.50	2.11
	319282	AA461358	Hs.12876	ESTs	1.00	1.00
	319289	W07304	Hs.79059	transforming growth factor, beta recepto	0.18	0.68
	319291	W85578	Hs.285243	hypothetical protein FLJ22029	0.26	0.62
	319293	F12119	Hs.12583	ESTs	3.13	4.50
10	319312	Z45481		gb:HSC2QE041 normalized infant brain cDN	1.10	1.00
	319370	H54254	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	0.16	0.73
	319391	R06304	Hs.13911	ESTs	1.26	2.43
	319396	H67130	Hs.301743	ESTs	0.70	0.76
	319398	AA359754	Hs.191196	ESTs	2.45	3.59
15	319407	R05329		gb:ye91b04.r1 Soares fetal liver spleen	2.00	3.54
	319425	T82930		gb:yd39i07.r1 Soares fetal liver spleen	4.28	8.81
	319433	R06050	Hs.191198	ESTs	6.15	14.13
	319437	AA282420	Hs.111991	ESTs, Weakly similar to Y48A5A.1 [C.eleg	3.26	5.68
	319466	AU809937	Hs.116417	ESTs	1.76	5.65
20	319471	R06546	Hs.191717	ESTs	4.29	4.84
	319480	R06933	Hs.184221	ESTs	1.00	1.00
	319484	T91772		gb:yd52a10.s1 Soares fetal liver spleen	2.81	4.88
	319486	AI382429	Hs.250799	ESTs	2.08	2.82
	319508	T99898	Hs.270104	ESTs, Moderately similar to ALU8_HUMAN A	2.80	4.39
	319523	T69499	Hs.191184	ESTs	1.55	3.25
25	319545	R83716	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	1.65	1.19
	319546	R09692		gb:yd23b12.r1 Soares fetal liver spleen	5.11	8.54
	319552	AA096106	Hs.20403	ESTs	1.89	3.36
	319582	T82998	Hs.250154	hypothetical protein FLJ12973	3.48	4.82
30	319586	D78808	Hs.283683	chromosome 8 open reading frame 4	0.26	0.82
	319604	R11679	Hs.297753	vimentin	1.68	3.41
	319609	AW247514	Hs.12293	hypothetical protein FLJ21103	3.06	4.24
	319611	H14957		gb:ym19c10.r1 Soares infant brain 1N1B H	2.76	4.24
	319653	AA770183	Hs.173515	uncharacterized hypothalamus protein HTO	2.61	3.55
35	319657	R19897	Hs.106604	ESTs	5.32	7.68
	319658	R13432	Hs.167481	synthrophin, gamma 1	3.35	5.00
	319661	H08035	Hs.21398	ESTs, Moderately similar to A Chain A, H	5.18	12.55
	319662	H06382	Hs.21400	ESTs	1.58	1.58
	319708	R15372	Hs.22664	ESTs	1.00	1.22
40	319742	T77668	Hs.21162	ESTs	2.48	3.13
	319748	R18178	Hs.295866	Homo sapiens mRNA; cDNA DKFZp434N1923 (f	3.02	4.85
	319772	R76633	Hs.22646	ESTs	4.36	11.61
	319788	AA321932	Hs.117414	KIAA1320 protein	2.56	3.68
	319805	R82857	Hs.271350	likely ortholog of mouse polydrom	4.63	6.56
45	319812	N74880	Hs.264330	N-acylsphingosine amidohydrolase (acid c	0.63	1.32
	319834	AA071267		gb:zm61g01.r1 Stratagene fibroblast (937	0.30	0.94
	319878	T78517	Hs.13941	ESTs	3.99	6.44
	319882	AA258981	Hs.291392	ESTs	5.09	7.36
	319912	T77559	Hs.94109	Homo sapiens cDNA FLJ13634 fis, clone PL	3.24	3.21
50	319935	H79460	Hs.271722	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.40	9.42
	319944	T79248	Hs.133510	ESTs	3.31	5.39
	319947	AA160967	Hs.14479	Homo sapiens cDNA FLJ14199 fis, clone NT	2.80	4.95
	319952	H06350	Hs.135056	Human DNA sequence from clone RP5-850E9	1.81	1.57
	320007	AA336314		gb:EST40943 Endometrial tumor Homo sapie	3.42	6.29
55	320018	T83263		gb:yd40h09.r1 Soares fetal liver spleen	2.77	5.14
	320030	H63789	Hs.296288	ESTs, Weakly similar to KIAA0638 protein	4.10	6.69
	320032	AI699772	Hs.292664	ESTs, Weakly similar to A46010 X-linked	3.27	3.27
	320040	AA233671	Hs.87164	hypothetical protein FLJ14001	1.81	1.64
	320047	T86564	Hs.302256	EST	3.38	7.36
60	320063	AA074108	Hs.120844	FOXJ2 forkhead factor	5.90	16.73
	320096	H58138	Hs.117915	ESTs	2.08	4.47
	320099	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	1.00	1.00
	320112	T92107	Hs.188489	ESTs	2.27	2.06
	320140	H94179	Hs.119023	SMC2 (structural maintenance of chromoso	1.00	1.00
65	320188	AW419200	Hs.172318	ESTs	1.26	1.00
	320193	AA831259	Hs.17132	ESTs	2.58	6.23
	320195	R62203	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	2.85	4.53
	320199	R78659	Hs.29792	ESTs	0.40	0.94
	320203	AL049227	Hs.124776	Homo sapiens mRNA; cDNA DKFZp564N1116 (f	0.84	1.18
70	320219	AA327584	Hs.127011	tubulointerstitial nephritis antigen	1.00	1.17
	320220	AF054910	Hs.127111	teklin 2 (testicular)	0.18	1.09
	320225	AF058989	Hs.128231	G antigen, family B, 1 (prostate associa	5.26	13.75
	320231	H03139	Hs.24683	ESTs	1.59	1.93
	320260	NM_003608	Hs.131924	G protein-coupled receptor 65	1.38	4.56
75	320267	AL049337	Hs.132571	Homo sapiens mRNA; cDNA DKFZp564P016 (fr	1.00	1.92
	320268	H06019	Hs.151293	Homo sapiens cDNA FLJ10664 fis, clone NT	5.58	5.70
	320322	AF077374	Hs.139322	small proline-rich protein 3	1.41	1.01
	320325	AI167978	Hs.139851	caveolin 2	0.05	0.67
	320330	AF026004	Hs.141660	chloride channel 2	2.17	1.26
80	320339	H10807	Hs.281434	Homo sapiens cDNA FLJ14028 fis, clone HE	1.81	2.32
	320388	H16055	Hs.31286	ESTs	1.00	3.22
	320402	R22291	Hs.23358	Homo sapiens clone FLC0578 PRO2852 mRNA,	1.41	1.36
	320413	AA203711	Hs.173269	ESTs	2.31	3.61
	320432	R62786	Hs.124136	ESTs	11.25	20.78
85	320436	AA253352	Hs.293663	ESTs	2.22	3.49
	320438	W24548	Hs.5669	ESTs	3.53	8.14

	320448	AI240233	Hs.80887	v-yes-1 Yamaguchi sarcoma viral related	1.42	3.46
	320451	R26944	Hs.180777	Homo sapiens mRNA; cDNA DKFZp554M0264 (f	0.87	0.81
	320484	AA094436	Hs.296267	folliculin-like 1	0.65	1.18
5	320499	R32555	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	3.44	7.15
	320514	AB007978	Hs.158278	KIAA0509 protein	6.44	13.62
	320521	N31454	Hs.24743	hypothetical protein FLJ20171	1.48	1.04
	320526	AW374205	Hs.111314	ESTs	3.66	7.87
	320527	R34672	Hs.324522	ESTs	3.16	5.63
10	320536	AA331732	Hs.137224	ESTs	2.83	5.83
	320556	AF054177	Hs.14570	hypothetical protein FLJ22530	1.28	1.00
	320564	AF055209	Hs.159395	peptidylglycine alpha-amidating monooxyg	1.22	0.81
	320587	Z44524	Hs.167456	Homo sapiens mRNA full length insert cDN	1.84	2.44
	320635	R54159	Hs.80506	small nuclear ribonucleoprotein polypept	1.00	6.25
	320639	AA243258	Hs.7395	hypothetical protein FLJ23182	2.60	2.30
15	320648	N48521	Hs.26549	Homo sapiens mRNA for KIAA1708 protein,	1.00	1.53
	320651	AA489268	Hs.111334	ferritin, light polypeptide	0.14	0.79
	320664	AI904216	Hs.91251	hypothetical protein FLJ11198	5.02	8.84
	320676	AA132650	Hs.300511	ESTs	3.63	5.37
	320683	R59291	Hs.26638	ESTs, Weakly similar to unnamed protein	0.37	1.31
20	320689	AA334609	Hs.171829	ESTs, Weakly similar to A54849 collagen	1.27	1.02
	320696	AW135016	Hs.172780	ESTs	3.53	4.60
	320714	AI445591		gb:yr04a10.r1 Soares fetal liver spleen	1.06	0.85
	320727	U96044	Hs.181125	immunoglobulin lambda locus	1.35	1.49
25	320771	AI793266	Hs.117176	poly(A)-binding protein, nuclear 1	0.04	0.82
	320794	AA281993	Hs.91226	ESTs	2.96	4.33
	320822	AF100780	Hs.194679	WNT1 inducible signaling pathway protein	0.10	0.79
	320824	AF120274	Hs.194689	artemin	1.16	1.11
	320830	AJ132445	Hs.266416	claudin 14	1.06	1.75
30	320843	AA317372	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	1.36	1.47
	320849	D60031	Hs.34771	ESTs	5.30	7.49
	320853	AI473796	Hs.135904	ESTs	1.00	1.00
	320896	AB002155	Hs.271580	uropod 1B	5.90	2.55
	320921	R94038	Hs.199538	inhibin, beta C	2.20	1.17
	320927	AI205786	Hs.213923	ESTs	0.18	1.46
35	320957	AI878933	Hs.92023	core histone macroH2A.2	1.67	2.18
	320997	H22544		gb:yn69f11.r1 Soares adult brain N2b5HB5	3.26	3.62
	321045	W88483	Hs.293650	ESTs	2.25	4.55
	321046	H27794	Hs.269055	ESTs	2.69	4.25
40	321052	AW372884	Hs.240770	nuclear cap binding protein subunit 2, 2	2.14	2.56
	321059	AI092824	Hs.126465	ESTs	1.69	0.53
	321062	R87955	Hs.241411	Homo sapiens mRNA full length insert cDN	2.76	5.20
	321067	AF131782	Hs.241438	Homo sapiens clone 24941 mRNA sequence	4.79	7.41
	321102	AA018306		gb:ze40d08.r1 Soares retina N2b4HR Homo	1.79	4.27
45	321130	H43750	Hs.125494	ESTs	1.00	3.14
	321142	AI817933	Hs.298351	ASPL protein	8.73	15.36
	321155	AA336635	Hs.99598	hypothetical protein MGC5338	3.04	5.03
	321158	AA700289		gb:yu76f11.r1 Soares fetal liver spleen	4.62	8.39
	321170	N53742	Hs.172982	ESTs	2.21	4.46
50	321199	AW385512		gb:yy56d10.s1 Soares multiple sclerosis	5.69	8.01
	321206	H54178	Hs.226469	Homo sapiens cDNA FLJ12417 fis, clone MA	4.00	7.32
	321225	AL080073	Hs.251414	Homo sapiens mRNA; cDNA DKFZp554B1462 (f	4.17	4.63
	321236	AW371941	Hs.18192	Ser/Arg-related nuclear matrix protein (1.00	1.00
	321244	AF058654		gb:Homo sapiens isolate AN.1 immunoglobu	2.18	9.13
55	321270	R83560		gb:yy76c06.s1 Soares fetal liver spleen	3.80	5.26
	321317	AI937060	Hs.6298	KIAA1151 protein	1.81	1.65
	321318	AB033041	Hs.137507	KIAA1215 protein	1.00	1.00
	321325	AB033100	Hs.300646	KIAA protein (similar to mouse paladin)	0.44	0.93
	321342	AA127984	Hs.222024	transcription factor BMAL2	4.94	4.93
60	321356	R93443	Hs.271770	ESTs	3.10	4.66
	321418	AI739161	Hs.161075	ESTs	2.28	2.54
	321420	AI368667	Hs.132743	ESTs	1.13	0.97
	321430	U05890		gb:H.sapiens (DIG3) mRNA for immunoglobu	2.42	3.35
	321453	N50080	Hs.82845	Homo sapiens cDNA: FLJ21930 fis, clone H	1.60	3.11
65	321467	X13075		gb:Human 2a12 mRNA for kappa-immunoglobu	0.42	0.72
	321468	AA514198	Hs.38540	ESTs	2.46	6.50
	321491	H70665	Hs.292549	ESTs	1.00	1.25
	321498	AW295517	Hs.255436	ESTs	3.19	6.24
	321504	W02356	Hs.268980	ESTs	2.28	3.86
70	321510	AA703650	Hs.255748	ESTs	2.14	3.94
	321513	H84972	Hs.108551	ESTs	2.78	5.37
	321516	AI382803	Hs.159235	ESTs	3.06	7.19
	321565	AI525773	Hs.268514	hypothetical protein FLJ11342	4.89	7.82
	321577	H84260		gb:ys90g04.r1 Soares retina N2b5HR Homo	1.00	1.73
75	321581	AA019964	Hs.28803	ESTs	4.88	6.73
	321582	AA143755	Hs.21858	trinucleotide repeat containing 3	1.00	2.08
	321587	H95531		gb:ys76e02.r1 Soares retina N2b4HR Homo	2.26	4.52
	321626	AA295430	Hs.96322	hypothetical protein FLJ23560	1.95	3.83
	321628	H87064	Hs.161051	ESTs, Moderately similar to ALU6_HUMAN A	0.47	1.02
80	321642	AW085917	Hs.247084	ESTs	1.52	1.38
	321669	H95404	Hs.294110	ESTs	2.17	2.45
	321687	AA625149		gb:af70c12.r1 Soares NhHMPu_S1 Homo sapi	4.31	6.95
	321688	H97646	Hs.123158	Homo sapiens cDNA FLJ12830 fis, clone NT	2.82	3.28
	321693	AA700017	Hs.173737	ras-related C3 botulinum toxin substrate	0.51	1.08
85	321700	N55160	Hs.167260	ESTs	4.57	7.46
	321701	AW390923	Hs.42568	ESTs	1.00	1.00

	321709	N25847	Hs.108923	RAB38, member RAS oncogene family	1.00	1.00
	321710	N35682	Hs.259743	ESTs	2.97	5.26
	321775	A1694875	Hs.202312	Homo sapiens clone N11 NTera2D1 teratoca	1.00	1.00
5	321777	A1637993	Hs.202312	Homo sapiens clone N11 NTera2D1 teratoca	1.68	0.45
	321779	N42729	Hs.163835	ESTs	0.90	0.90
	321829	D81993	Hs.8966	tumor endothelial marker 8	2.69	3.89
	321846	AA281594	Hs.87902	ESTs	5.11	7.64
	321879	AL109670	Hs.302809	ESTs	6.49	9.58
	321883	AA426494	Hs.46901	KIAA1462 protein	0.28	0.95
10	321899	N55158	Hs.29468	ESTs	0.39	0.95
	321911	AF026944	Hs.293797	ESTs	6.20	10.76
	321949	R49202	Hs.181694	EST	4.62	10.51
	321955	A1651866	Hs.195689	ESTs	2.89	5.47
	321956	AL110177	Hs.132882	ESTs	0.32	1.25
15	321987	AL133612	Hs.272759	KIAA1457 protein	1.00	1.83
	321991	AL133627	Hs.158923	Homo sapiens mRNA; cDNA DKFZp434K0722 (f	4.00	6.47
	322002	AA328801	Hs.84522	ESTs	2.10	3.48
	322035	AL137517	Hs.306201	hypothetical protein DKFZp564O1278	1.00	1.90
20	322044	AW340926		gbxy51b10.x1 NCI_CGAP_Lu34.1 Homo sapie	3.20	9.67
	322057	N92197	Hs.154679	synaptotagmin 1	1.55	1.07
	322060	A1341937		gb:qt10e03.x1 NCI_CGAP_GC4 Homo sapiens	4.59	7.88
	322070	U80769	Hs.210322	Homo sapiens mRNA for KIAA1766 protein,	2.78	4.52
	322083	AF074982	Hs.226031	ESTs, Highly similar to KIAA0535 protein	3.10	5.52
	322091	A1819863	Hs.106243	ESTs	1.59	1.75
25	322125	R93901		gb:yq16c12.r1 Soares fetal liver spleen	2.06	5.27
	322130	R98978	Hs.117767	ESTs	10.12	16.49
	322147	AF085919	Hs.114176	ESTs	0.94	0.64
	322166	AF085958		gb:yr88b03.r1 Soares fetal liver spleen	4.09	6.67
30	322173	H52567		gb:yt85d04.r1 Soares_pineal_gland_N3HPG	3.46	4.85
	322178	H56535		gb:yt88g03.r1 Soares_pineal_gland_N3HPG	0.44	2.54
	322179	H92891		gb:yt94c02.s1 Soares_pineal_gland_N3HPG	4.52	7.50
	322186	H57346	Hs.269187	ESTs	0.15	0.98
	322196	W87895	Hs.211516	ESTs	2.20	5.04
35	322212	AF087995	Hs.134877	ESTs	3.42	4.84
	322221	A1890619	Hs.179662	nucleosome assembly protein 1-like 1	0.82	2.14
	322277	A1640193	Hs.226389	ESTs	3.62	3.98
	322278	AF086283		gb:zd46f01.r1 Soares_fetal_heart_NbHH19W	1.00	1.00
40	322284	A1792140	Hs.49265	ESTs	0.66	2.76
	322288	AL037273	Hs.7886	pellino (Drosophila) homolog 1	0.71	0.70
	322320	AF086419		gb:zd78d03.r1 Soares_fetal_heart_NbHH19W	2.02	2.76
	322336	AA308526	Hs.76152	decorin	2.92	4.44
	322339	W17348		gb:zb18c07.x5 Soares_fetal_lung_NbHL19W	8.50	11.56
	322366	AW404274	Hs.122492	hypothetical protein	0.61	1.34
45	322372	W25624	Hs.153943	ESTs	7.37	12.07
	322374	A1394663	Hs.122116	ESTs, Moderately similar to Ost2 [M.musc	4.78	10.50
	322378	AF064819	Hs.201877	DESC1 protein	1.00	1.00
	322388	A1815730	Hs.247474	hypothetical protein FLJ21032	7.09	8.49
	322416	AA223183	Hs.298442	adaptor-related protein complex 3, mu 1	3.20	5.80
50	322419	AA248997	Hs.14084	ring finger protein 7	1.64	1.57
	322425	W37943	Hs.34892	KIAA1323 protein	0.83	1.00
	322431	AA069222	Hs.141892	ESTs	3.96	5.22
	322450	AA040131	Hs.25144	ESTs	5.18	12.67
	322465	AA137152	Hs.265049	phosphoserine aminotransferase	3.41	2.23
55	322467	AF116826	Hs.180340	putative protein-tyrosine kinase	1.00	1.30
	322473	AA744286	Hs.266935	tRNA selenocysteine associated protein	1.75	2.03
	322509	T52172	Hs.302213	ESTs	1.00	2.27
	322523	W80398	Hs.193197	ESTs	2.75	5.49
	322527	AF147359		gb:Homo sapiens full length insert cDNA	1.25	1.27
60	322560	A1916847	Hs.270947	ESTs	4.57	8.81
	322566	W87285	Hs.269587	ESTs	1.00	1.42
	322585	AA837622		gb:zh69c01.r1 Soares_fetal_liver_spleen_	4.18	6.94
	322635	AA679084		gb:zh90h08.r1 Soares_fetal_liver_spleen_	2.40	4.85
	322641	AA007352	Hs.256042	ESTs	2.94	4.64
65	322653	A1828854	Hs.258538	striatin, calmodulin-binding protein.	0.48	0.38
	322664	AA011522		gb:zi03g07.r1 Soares_fetal_liver_spleen_	1.92	2.18
	322687	A1110759		gb:AF074666 Human fetal liver cDNA libra	4.14	6.75
	322692	AA018117	Hs.60843	potassium voltage-gated channel, shaker-	3.50	5.00
	322694	A1110872	Hs.279812	PRO0327 protein	1.80	1.72
70	322708	AF113674	Hs.283773	clone FLB1727	1.00	3.43
	322712	AA021328	Hs.23507	hypothetical protein FLJ11109	3.28	3.86
	322766	AW058805	Hs.288457	Homo sapiens cDNA FLJ12280 fis, clone MA	1.63	1.53
	322770	AA045796	Hs.122682	ESTs	1.53	1.05
	322794	A1608591	Hs.38991	S100 calcium-binding protein A2	12.05	1.94
75	322810	A1962276	Hs.127444	ESTs	4.09	6.90
	322818	AW043782	Hs.293616	ESTs	1.20	1.63
	322820	A1377755	Hs.120695	ESTs	0.21	1.93
	322872	AA827228	Hs.126943	ESTs	2.04	1.63
	322882	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	5.28	1.22
80	322887	A1986306	Hs.86149	phosphoinositol 3-phosphate-binding prot	2.80	2.24
	322913	A1733737	Hs.68837	ESTs	2.38	6.61
	322926	A1825940	Hs.211192	ESTs	4.02	5.79
	322929	A1365585	Hs.146246	ESTs	0.30	1.14
	322958	A1905228	Hs.83484	SRY (sex determining region Y)-box 4	2.06	1.13
85	322971	C15953	Hs.212760	hypothetical protein FLJ13649	1.18	2.00
	322981	AA493252	Hs.159577	ESTs	2.28	2.61

	322988	C18727	Hs.171941	ESTs	0.39	2.00
	323003	A1733859	Hs.149089	ESTs	3.28	1.00
	323013	AA134042	Hs.191451	ESTs	3.38	5.68
5	323025	AL157565	Hs.315369	Homo sapiens cDNA: FLJ23075 fis, clone L	0.06	1.10
	323032	AW244073	Hs.145946	ESTs	10.18	21.27
	323052	R21124	Hs.85573	Homo sapiens DC29 mRNA, complete cds	1.46	1.90
	323064	AL119341	Hs.49359	Homo sapiens mRNA; cDNA DKFZp547E052 (fr	3.08	5.64
	323098	A1700025	Hs.270471	ESTs	2.31	4.49
10	323102	AL119913	Hs.163615	ESTs	5.38	11.64
	323155	AL135041		gb:DKFZp762K2310_r1 762 (synonym: hmel2)	2.38	5.56
	323176	AW071648	Hs.82101	pleckstrin homology-like domain, family	1.06	1.41
	323191	AA195600	Hs.301570	ESTs	0.73	1.24
	323225	AA205654	Hs.24790	KIAA1573 protein	5.25	11.95
15	323232	AA148722	Hs.224680	ESTs	0.45	1.35
	323256	AW003382	Hs.243886	nuclear autoantigenic sperm protein (his	1.71	1.83
	323281	A1697555	Hs.292659	ESTs	1.24	3.21
	323283	AA256014	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	12.68	15.05
	323314	AA226310	Hs.191501	ESTs	4.42	9.61
20	323316	AL134620	Hs.280175	ESTs	2.98	5.93
	323334	A1336501	Hs.77273	ras homolog gene family, member A	1.98	3.30
	323338	R74219	Hs.23348	S-phase kinase-associated protein 2 (p45	1.62	1.00
	323348	AA233056	Hs.191518	ESTs	1.00	1.07
	323351	AA704103	Hs.24049	ESTs	1.43	1.68
25	323359	AA234172	Hs.137418	ESTs	0.34	1.18
	323360	AA716061	Hs.161719	ESTs	3.01	3.71
	323405	AW139550	Hs.115173	ESTs	1.90	8.81
	323420	A1672386	Hs.263760	ESTs	0.29	1.01
	323434	AW081455	Hs.120219	ESTs	2.27	1.92
30	323445	AA253103	Hs.135569	ESTs, Weakly similar to NEUROD (Hsapien	0.43	0.80
	323449	AA232865	Hs.284153	Fanconi anemia, complementation group A	3.19	3.85
	323492	H00978	Hs.20887	hypothetical protein FLJ10392	2.70	3.20
	323501	AA182461	Hs.84520	ESTs	2.04	3.31
	323505	A1652287		gb:EST382593 MAGE resequences, MAGK Homo2.21		3.08
35	323515	AA282274	Hs.256083	ESTs	2.69	3.40
	323541	A1185116	Hs.104813	RP42 homolog	1.20	1.09
	323545	A1814405	Hs.224569	ESTs	1.25	1.55
	323635	R63117	Hs.9691	Homo sapiens cDNA: FLJ23249 fis, clone C	0.27	0.72
40	323675	AA984759	Hs.272168	tumor differentially expressed 1	3.70	5.80
	323678	AL042121	Hs.20880	ESTs	3.33	5.10
	323691	AA317561	Hs.145599	ESTs	1.00	1.00
	323693	AW297758	Hs.249721	ESTs	2.01	1.54
	323746	AW298611	Hs.12808	MARK	4.11	5.53
	323774	AA329806	Hs.321056	Homo sapiens mRNA; cDNA DKFZp586F1322 (f	2.06	3.70
45	323856	AA355264	Hs.267604	hypothetical protein FLJ10450	3.42	8.13
	323857	T18988	Hs.293668	ESTs	5.97	12.51
	323870	AA341774	Hs.129212	ESTs	3.17	4.52
	323876	AL042492	Hs.147313	ESTs	0.36	1.00
	323885	AA344308	Hs.128427	Homo sapiens BAC clone RP11-335J18 from	2.31	3.33
50	323911	AL043212	Hs.92550	ESTs	4.38	5.41
	323919	AA862973	Hs.220704	ESTs	5.80	10.20
	323972	A1869964	Hs.182906	ESTs	3.10	5.14
	324005	AA610011	Hs.208021	ESTs	5.34	10.07
	324036	A1472078	Hs.303662	ESTs	1.00	5.03
55	324055	AA528794	Hs.128644	ESTs	0.86	1.00
	324063	AW292740	Hs.272813	dual oxidase 1	0.45	0.91
	324072	AA381829		gb:EST94855 Activated T-cells I Homo sap	2.82	5.12
	324092	AW269931	Hs.202473	Homo sapiens cDNA: FLJ22278 fis, clone H	2.40	2.52
	324095	AW377983	Hs.298140	Homo sapiens cDNA: FLJ22502 fis, clone H	1.32	4.30
60	324129	A1381918	Hs.285833	Homo sapiens cDNA: FLJ22135 fis, clone H	1.40	1.77
	324132	AW504860	Hs.288836	hypothetical protein FLJ12673	4.24	6.21
	324214	AA412395	Hs.225740	ESTs	6.96	10.69
	324227	AA295552	Hs.28631	Homo sapiens cDNA: FLJ22141 fis, clone H	0.81	0.53
	324266	AL047634	Hs.231913	ESTs	2.42	4.05
	324275	AA429088	Hs.98523	ESTs	3.62	5.38
65	324281	AL048026	Hs.124675	ESTs, Weakly similar to T14742 hypotheti	0.14	0.70
	324290	AA432032	Hs.304420	ESTs	3.71	4.34
	324303	AL118754		gb:DKFZp761P1910_r1 761 (synonym: hamy2)	0.95	0.91
	324312	A1188841	Hs.128173	ESTs	4.06	5.91
70	324325	AL138153	Hs.300410	ESTs	5.88	8.25
	324338	AL138357	Hs.145078	regulator of differentiation (In S. pombe	0.87	1.25
	324341	AW197734	Hs.99807	ESTs, Weakly similar to unnamed protein	1.28	1.00
	324343	AW452016	Hs.293232	ESTs	2.54	3.46
	324371	AA452305	Hs.270319	ESTs	5.85	8.36
75	324382	AW502749	Hs.24724	MFH-amplified sequences with leucine-ric	0.76	1.64
	324384	AA453396	Hs.127656	KIAA1349 protein	2.88	5.69
	324385	F28212	Hs.284247	KIAA1491 protein	1.81	1.99
	324388	A1924963	Hs.306206	hypothetical protein FLJ11215	1.00	1.00
	324432	AA464510	Hs.152812	ESTs	2.73	2.17
80	324497	AW152624	Hs.136340	ESTs, Weakly similar to unnamed protein	0.71	1.90
	324510	A1148353	Hs.287425	Homo sapiens cDNA FLJ11569 fis, clone HE	1.00	1.00
	324580	AA492588		gb:ng99c08.s1 NCI_CGAP_Thy1 Homo sapiens	2.18	3.50
	324582	AA506935	Hs.132036	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.96	11.36
	324633	AA572984	Hs.325489	ESTs	2.92	4.22
	324640	AW295332	Hs.134798	ESTs, Moderately similar to TTL MOUSE TU	5.48	11.74
85	324675	AW014734	Hs.157969	ESTs	0.39	0.73

	324699	AW504732	Hs.21275	hypothetical protein FLJ11011	0.93	0.93
	324747	AA603532	Hs.130807	ESTs	1.57	1.81
	324748	AA657457	Hs.292385	ESTs	1.55	1.34
5	324801	AI819924	Hs.14553	sterol O-acyltransferase (acyl-Coenzyme	1.00	6.56
	324804	AI692552		gb:wd73f12.x1 NCI_CGAP_Lu24 Homo sapiens	1.00	7.53
	324828	AA843926	Hs.124434	ESTs	2.00	3.25
	324855	AW152305	Hs.122364	ESTs	2.74	3.43
	324866	AI541214	Hs.46320	Small proline-rich protein SPRK (human,	1.07	0.95
10	324871	AW297755	Hs.271923	Homo sapiens cDNA: FLJ22785 fis, clone K	1.68	1.21
	324886	AA806794	Hs.131511	ESTs	2.56	5.61
	324889	D31010		gb:HUM12147 Human fetal lung Homo sapie	2.20	4.65
	324948	AW383818	Hs.265459	ESTs, Moderately similar to ALU2_HUMAN A	5.28	7.05
	324953	AI264628	Hs.125428	ESTs	3.37	5.51
15	324958	AA625076	Hs.132892	protocadherin 20	5.12	9.81
	324988	T06997	Hs.121028	hypothetical protein FLJ10549	2.52	1.08
	325024	F13254	Hs.78572	laminin, alpha 4	5.24	10.22
	325105	H97109	Hs.105421	ESTs	1.00	1.00
	325108	AA401863	Hs.22380	ESTs	1.99	2.14
20	325114	D83901	Hs.315562	ESTs	2.73	3.17
	325146	AI064690	Hs.171176	ESTs	1.86	3.41
	325149	D61117	Hs.187646	ESTs	0.42	0.93
	325187	AI653682	Hs.197812	ESTs	6.50	11.31
	325228				6.18	15.76
25	325235				2.64	4.12
	325328				2.87	4.42
	325340				0.29	0.33
	325367				16.56	24.29
30	325373				0.63	1.22
	325389				0.88	1.05
	325436				5.75	14.14
	325471				8.46	17.82
	325498				3.32	6.42
	325557				5.51	8.28
35	325559				7.48	21.40
	325560				4.08	6.25
	325569				4.20	5.24
	325585				1.10	1.13
	325587				1.00	1.00
40	325597				2.98	13.40
	325639				0.78	0.78
	325685				0.48	0.66
	325686				0.95	1.55
	325735				4.48	9.20
45	325739				0.69	0.88
	325740				2.42	6.61
	325792				7.88	9.83
	325819				4.74	7.18
	325883				2.02	2.64
50	325895				7.78	15.98
	325925				2.04	10.60
	325932				4.18	7.36
	325941				3.66	9.03
	325969				0.61	0.80
55	325971				4.88	7.42
	326025				0.55	1.07
	326046				7.21	14.72
	326099				3.60	5.98
	326108				1.27	1.06
60	326163				3.27	5.70
	326165				0.45	1.11
	326189				0.13	0.45
	326204				5.60	9.00
	326230				7.00	12.01
	326274				1.00	8.09
65	326360				9.86	15.35
	326393				0.52	0.77
	326505				1.00	1.42
	326515				1.24	5.84
70	326589				9.20	13.49
	326592				2.77	4.01
	326605				2.01	2.53
	326692				1.00	1.00
	326693				1.00	1.31
75	326720				0.19	0.65
	326742				2.34	7.20
	326770				0.25	0.83
	326818				3.09	4.56
	326936				2.08	3.45
80	326964				0.41	1.70
	326983				2.02	3.80
	326991				1.09	1.20
	327036				1.00	8.04
	327040				3.05	4.22
85	327053				3.55	6.31
	327075				1.59	1.40

	327085	2.50	12.57
	327130	5.38	8.04
	327156	3.74	6.58
5	327220	1.28	1.54
	327224	6.56	12.91
	327288	2.61	5.40
	327321	2.42	3.11
	327332	6.62	10.58
	327361	2.69	4.41
10	327377	2.04	6.72
	327395	2.61	4.50
	327414	1.00	8.01
	327442	5.91	9.65
	327467	6.58	18.01
15	327473	3.79	7.48
	327483	4.08	8.87
	327562	0.68	2.86
	327568	1.00	2.00
	327606	2.06	3.61
20	327611	5.90	14.26
	327642	4.06	8.74
	327654	1.05	2.08
	327734	1.00	1.00
	327775	1.46	11.79
25	327796	3.47	5.65
	327840	3.26	6.64
	327940	5.84	15.58
	327984	0.36	1.50
	328004	1.87	1.42
30	328021	0.42	0.59
	328058	2.83	4.68
	328100	3.04	5.39
	328101	3.54	5.20
	328113	0.72	0.91
35	328157	5.58	5.16
	328196	5.76	11.13
	328197	5.98	10.58
	328264	3.11	4.88
	328299	2.20	3.06
40	328342	1.49	1.94
	328365	1.00	1.00
	328369	4.40	7.36
	328381	1.86	4.93
45	328451	5.51	7.56
	328481	0.13	0.72
	328500	2.71	3.97
	328530	5.41	7.62
	328600	3.14	10.68
50	328608	4.56	8.17
	328616	2.24	11.91
	328623	3.04	5.46
	328632	0.70	1.19
	328664	3.48	6.80
55	328666	10.42	26.47
	328698	9.68	14.56
	328700	2.74	10.22
	328708	0.15	0.57
	328735	6.23	8.91
60	328743	3.62	6.54
	328806	0.22	0.78
	328861	3.68	10.54
	328908	5.42	16.36
	328933	2.02	5.29
65	328934	1.73	4.45
	328949	3.34	5.41
	329005	2.88	7.26
	329011	2.52	3.72
	329033	1.00	1.03
70	329037	5.07	8.16
	329067	1.98	2.41
	329134	2.24	3.25
	329157	2.30	11.04
	329178	2.64	5.02
75	329192	6.41	15.27
	329194	0.31	0.79
	329204	1.60	3.75
	329224	2.99	6.11
	329228	0.83	0.83
80	329288	0.63	1.01
	329337	1.00	1.00
	329541	0.76	1.68
	329560	1.34	2.02
	329588	1.68	2.22
85	329643	4.18	11.77
	329703	1.00	1.00

	329764			5.78	15.50
	329816			2.09	5.44
	329860			3.13	10.77
	329993			7.83	14.21
5	330020			5.58	13.12
	330036			3.32	5.57
	330052			4.31	7.97
	330085			1.34	1.76
	330088			4.70	12.46
10	330093			0.44	1.06
	330100			3.47	4.83
	330106			2.14	3.61
	330107			3.17	6.87
	330120			5.51	11.89
15	330123			4.50	12.74
	330208			1.55	7.62
	330263			13.10	23.38
	330300			2.81	4.98
	330313			3.00	4.41
20	330366			0.67	0.76
	330372			4.76	11.82
	330385	AA449749	Hs.182971	2.14	2.15
	330397	D14659	Hs.154387	0.40	1.15
25	330468	L10343	Hs.112341	1.11	0.94
	330472	L24203	Hs.82237	1.67	1.17
	330478	L39486	Hs.295049	0.46	1.07
	330493	M27826	Hs.267319	1.07	0.95
	330495	M31328	Hs.71642	0.97	0.96
	330506	M61906	Hs.6241	0.17	3.66
30	330512	M80563	Hs.81256	0.60	1.06
	330537	U19765	Hs.2110	2.81	2.07
	330547	U32989	Hs.183671	3.91	1.49
	330551	U39840	Hs.299867	1.15	1.03
	330568	U56244		2.83	4.79
35	330599	U90437		2.08	1.54
	330601	U90916	Hs.82845	0.89	1.35
	330605	X02419	Hs.77274	1.87	1.55
	330609	X04741	Hs.76118	1.83	1.30
	330617	X53587	Hs.85266	1.54	1.15
40	330630	X78669	Hs.79088	1.39	1.19
	330644	Y07755	Hs.38991	3.83	1.13
	330650	Z68228	Hs.2340	1.25	0.95
	330660	AA347868	Hs.139293	15.50	29.07
	330692	AA017045	Hs.6702	1.00	1.00
45	330707	AA133891	Hs.293690	0.20	1.35
	330715	AA233707	Hs.11571	0.12	1.40
	330717	AA233926	Hs.52620	6.62	5.42
	330722	AA243560	Hs.34382	1.40	1.65
	330740	AA297746	Hs.22654	0.27	2.04
50	330742	AA400979	Hs.25691	0.44	0.90
	330744	AA406142	Hs.12393	0.71	3.23
	330751	AA428286	Hs.29543	1.66	1.52
	330760	AA448663	Hs.30469	0.52	0.90
	330763	AA450200	Hs.274337	0.37	0.97
55	330786	D60374	Hs.49136	0.78	0.84
	330790	T48536	Hs.105807	0.23	3.17
	330814	AA015730	Hs.265398	0.37	2.07
	330827	AA040332	Hs.12744	1.60	1.00
60	330844	AA063037	Hs.65803	0.93	1.16
	330901	AA157818	Hs.267319	1.02	1.03
	330931	F01443	Hs.284256	0.24	0.88
	330952	H02855	Hs.29567	0.08	1.31
	330961	H10998	Hs.7164	1.29	1.26
65	330968	H16568	Hs.23748	0.48	0.96
	331014	H98597	Hs.30340	0.29	0.74
	331046	N65653	Hs.191358	0.99	8.56
	331060	N75081	Hs.157148	1.24	1.00
	331099	R36671	Hs.83937	0.75	1.03
70	331108	R41408	Hs.21983	1.00	2.75
	331131	R54797		6.04	10.68
	331135	R61398	Hs.4197	0.80	0.96
	331170	T23461	Hs.159293	2.63	4.29
	331180	T32446	Hs.6640	1.78	2.71
	331183	T40769	Hs.8469	1.00	3.01
75	331203	T82310		1.70	3.80
	331271	AA059347	Hs.82226	1.20	3.19
	331306	AA252079	Hs.63931	0.31	1.30
	331327	AA281076	Hs.109221	2.09	2.41
	331341	AA303125	Hs.23240	0.72	2.43
80	331359	AA416979	Hs.46901	0.09	0.91
	331363	AA421562	Hs.91011	1.02	0.87
	331378	AA448881	Hs.49282	1.03	1.23
	331384	AA456001	Hs.93847	1.40	1.00
	331402	AA505135	Hs.44037	1.80	3.93
85	331422	F10802	Hs.163628	1.65	1.89

	331490	N32912	Hs.26813	CDA14	2.48	1.73
	331531	N51343		gbxyz15g04.s1 Soares_multiple_sclerosis_	0.98	1.68
	331547	N54811		gbxod74f04.s1 NCL_CGAP_Ov2 Homo sapiens	3.80	5.75
5	331578	N67960	Hs.249989	ESTs	0.11	0.67
	331589	N71027	Hs.152618	ESTs	1.09	1.38
	331608	N89861	Hs.112110	PTD007 protein	0.93	0.76
	331614	N92293	Hs.240272	EST	0.17	1.34
	331668	W69707	Hs.58030	EST	2.24	3.82
10	331671	W72033	Hs.194695	ras homolog gene family, member 1	1.00	1.24
	331676	W79834	Hs.58559	ESTs, Weakly similar to rholekin (Musc	0.08	1.07
	331681	W85712	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	8.72	4.27
	331692	W93592	Hs.152213	wingless-type MMTV integration site fami	0.94	0.54
	331717	AA190888	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par	1.57	1.34
15	331718	AA191404	Hs.104072	ESTs	6.80	11.77
	331811	AA404500	Hs.301570	ESTs	1.10	1.00
	331820	AA405970	Hs.97996	transcription termination factor, mitoc	0.73	0.59
	331831	AA412031	Hs.97901	EST	2.77	4.08
	331852	AA418988	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	0.23	0.93
20	331943	AA453418	Hs.21275	hypothetical protein FLJ11011	0.36	1.88
	331969	AA460702	Hs.82772	collagen, type XI, alpha 1	1.00	1.00
	331990	AA478102	Hs.139631	ESTs	3.04	3.87
	332002	AA482009	Hs.105104	ESTs	1.19	0.78
	332027	AA489671	Hs.65541	hypothetical protein FLJ20073	1.27	1.03
25	332029	AA489697	Hs.145053	ESTs	0.30	1.82
	332033	AA489840	Hs.251014	EST	2.30	3.70
	332048	AA496019	Hs.201591	ESTs	0.17	0.52
	332071	AA598594	Hs.205293	KIAA1211 protein	1.35	1.23
	332074	AA599012		gb:ae41e11.s1 Gessler Wilms tumor Homo s	0.19	2.00
30	332083	AA600200	Hs.155546	KIAA1080 protein; Golgi-associated, gamm	0.31	1.18
	332085	AA600353	Hs.173933	nuclear factor I/A	0.30	1.50
	332125	AA609861	Hs.312447	ESTs	0.22	0.62
	332177	F10812	Hs.101433	ESTs	8.21	18.03
	332180	H03348	Hs.7327	claudin 1	2.27	1.57
35	332185	H10356	Hs.101689	ESTs	0.09	1.18
	332203	H49388	Hs.317769	EST	8.05	5.02
	332232	N48891	Hs.101915	Stargardt disease 3 (autosomal dominant)	0.78	0.85
	332240	N54803	Hs.324267	ESTs, Weakly similar to putative p150 [0.96	1.23
	332261	N70294	Hs.269137	ESTs	2.40	3.74
40	332275	R08838	Hs.26530	serum deprivation response (phosphatidyl	0.27	0.75
	332280	R38100	Hs.146381	RNA binding motif protein, X chromosome	0.39	1.88
	332299	R69250	Hs.21201	nectin 3; DKFZP566B0846 protein	5.24	12.76
	332304	R74041	Hs.101539	ESTs	1.44	3.18
	332314	T25862	Hs.101774	hypothetical protein FLJ23045	0.68	1.32
45	332384	M11433	Hs.101850	retinol-binding protein 1, cellular	1.71	0.88
	332434	N75542	Hs.289068	Homo sapiens cDNA FLJ11918 fis, clone HE	0.43	0.86
	332445	T63781	Hs.11112	ESTs	0.68	1.00
	332453	L00205	Hs.111758	keratin 6A	31.54	1.00
	332458	M33493	Hs.250700	trypsin beta 1	0.51	1.00
50	332504	AA053917	Hs.15106	chromosome 14 open reading frame 1	0.79	1.24
	332525	M17252	Hs.278430	cytochrome P450, subfamily XXA (steroid	0.98	1.70
	332530	M31682	Hs.1735	inhibin, beta B (activin AB beta polypep	0.88	0.66
	332535	N20284	Hs.19280	cysteine-rich motor neuron 1	0.22	1.46
	332539	AA412528	Hs.20183	ESTs, Weakly similar to AF164793 1 prote	0.93	1.49
55	332559	M13955	Hs.166189	cytokeratin 2	0.35	1.13
	332563	N92924	Hs.274407	protease, serine, 16 (thymus)	1.00	1.00
	332565	AA234896	Hs.25272	E1A binding protein p300	0.36	1.05
	332594	AA279313	Hs.3239	methyl CpG binding protein 2 (Rett syndr	0.53	0.59
	332634	S38953	Hs.283750	tenascin XA	0.38	1.16
60	332638	AA283034	Hs.50640	JAK binding protein	1.00	1.70
	332640	AA417152	Hs.5101	protein regulator of cytokinesis 1	6.15	1.16
	332654	AA001296	Hs.288217	hypothetical protein MGC2941	1.50	2.73
	332665	AA223335	Hs.63788	propionyl Coenzyme A carboxylase, beta p	1.20	0.91
	332692	AA496035	Hs.247926	gap junction protein, alpha 5, 40kD (con	0.17	1.12
65	332716	L00058	Hs.79070	v-myc avian myelocytomatosis viral oncog	1.00	1.44
	332736	L13773	Hs.114765	myeloid/lymphoid or mixed-lineage leukem	1.00	1.81
	332758	X93921	Hs.286938	dual specificity phosphatase 7	0.53	0.78
	332781	AA233258	Hs.247112	hypothetical protein FLJ10902	1.44	1.56
	332792				1.70	1.19
70	332816				1.85	2.47
	332858				1.04	1.57
	332906				3.48	8.04
	332911				1.00	1.00
	332912				1.06	4.40
75	332922				1.00	1.00
	332956				0.42	0.88
	332959				1.96	6.34
	332962				0.56	0.99
	332984				0.30	0.78
80	332998				1.47	2.01
	333058				0.47	1.38
	333097				2.14	3.19
	333121				2.76	3.70
	333122				1.92	1.21
85	333123				1.85	1.39
	333138				0.47	0.52

	333139	1.88	0.84
	333140	0.21	0.64
	333221	1.51	1.11
5	333260	0.75	1.01
	333380	6.68	15.75
	333387	4.56	12.61
	333512	5.05	8.01
	333524	2.28	3.98
10	333585	2.31	1.53
	333603	2.23	1.17
	333604	2.51	1.58
	333618	0.52	0.98
	333627	1.44	1.36
15	333628	1.90	1.90
	333650	1.85	2.10
	333678	1.85	2.35
	333750	2.18	5.67
	333763	1.99	2.60
20	333767	1.02	0.96
	333768	1.78	1.65
	333769	2.15	2.13
	333772	1.46	2.53
	333777	1.00	1.42
25	333846	2.99	4.50
	333894	0.47	0.94
	333887	0.50	1.00
	333891	0.43	0.89
	333892	0.51	0.91
30	333904	0.26	1.13
	333906	0.55	0.98
	333948	1.70	2.15
	333954	0.37	1.09
	333966	8.10	14.30
35	333968	0.63	1.38
	334061	4.24	12.30
	334094	1.30	12.03
	334113	4.55	8.63
	334161	0.82	1.59
40	334183	0.47	0.76
	334187	1.36	3.70
	334219	0.69	1.04
	334222	1.88	1.70
	334223	4.72	3.14
45	334239	0.79	0.62
	334255	0.45	1.10
	334333	1.00	3.56
	334378	3.98	5.76
	334382	1.50	1.31
50	334492	3.59	4.75
	334562	5.94	15.40
	334588	8.14	19.53
	334616	1.55	1.56
	334633	5.16	8.07
55	334648	0.59	2.13
	334787	3.70	7.15
	334866	8.13	10.60
	334891	0.32	1.14
	334933	1.00	3.84
60	334934	4.01	7.43
	334945	1.04	2.96
	334967	0.29	1.14
	334990	1.50	1.39
	335015	5.88	18.65
65	335093	0.55	1.75
	335120	4.31	8.01
	335125	0.38	1.97
	335179	1.24	1.98
	335188	0.46	1.47
70	335211	1.61	1.42
	335288	0.73	0.97
	335289	0.20	0.26
	335361	2.18	1.58
	335379	0.50	0.71
75	335414	3.64	14.94
	335416	2.93	3.98
	335496	0.96	0.91
	335497	1.71	1.92
	335548	1.15	2.40
80	335551	3.22	10.54
	335558	3.42	4.89
	335586	5.50	12.75
	335619	2.99	3.07
	335620	3.80	8.29
85	335621	0.28	0.57
	335682	0.46	1.17

	335686	2.55	3.81
	335755	2.24	1.07
	335784	0.20	0.97
5	335814	1.13	1.48
	335815	2.45	3.51
	335823	1.00	4.16
	335835	0.49	1.70
	335851	1.65	1.39
10	335868	2.98	6.43
	335896	0.98	0.99
	335936	12.10	21.93
	335948	1.00	1.64
	335983	1.00	4.21
	335995	0.37	1.17
15	336021	1.04	0.84
	336034	11.40	23.54
	336038	1.19	1.21
	336066	0.54	1.63
20	336107	0.95	0.70
	336205	3.13	6.29
	336275	3.20	10.10
	336292	2.34	3.09
	336331	1.00	1.00
25	336419	0.65	0.79
	336632	2.33	2.16
	336633	2.55	2.23
	336634	2.19	2.03
	336635	2.69	2.48
30	336636	2.13	1.83
	336637	2.43	2.24
	336638	2.31	2.03
	336659	0.60	1.31
	336675	0.31	1.18
35	336684	1.50	1.14
	336694	4.74	7.10
	336716	4.43	6.37
	336721	2.20	0.74
	336738	1.64	2.14
40	336900	6.14	12.73
	336948	1.00	1.00
	337028	1.30	2.09
	337043	4.01	11.53
	337046	1.67	1.84
45	337054	2.78	7.35
	337128	7.20	16.14
	337162	3.45	5.34
	337183	5.72	11.41
	337184	3.72	5.90
50	337192	1.27	1.05
	337194	1.88	1.68
	337229	0.22	1.03
	337268	1.00	3.31
	337299	3.23	5.14
55	337325	2.76	3.72
	337389	5.80	10.42
	337493	2.06	6.30
	337497	7.88	20.29
	337500	3.80	4.48
60	337549	1.66	2.31
	337603	1.27	8.54
	337605	5.76	7.16
	337671	0.73	0.97
	337755	1.54	0.92
65	337786	5.07	9.73
	337809	6.18	12.87
	337862	3.78	12.97
	337871	2.66	8.16
	337958	0.26	1.34
70	338008	1.48	1.12
	338033	2.38	14.59
	338083	0.65	2.16
	338110	1.00	1.61
	338112	5.86	8.25
75	338145	1.70	1.97
	338148	8.07	18.19
	338158	1.30	4.55
	338161	2.58	3.57
	338179	1.00	1.00
80	338182	3.32	4.63
	338189	1.00	3.34
	338197	0.99	1.69
	338199	4.58	7.62
	338215	6.01	15.85
85	338279	0.63	0.95
	338316	20.58	38.66

	338322	3.23	7.39
	338357	4.10	11.39
	338359	10.12	21.59
5	338366	0.69	1.02
	338374	0.40	1.18
	338414	0.47	1.06
	338418	6.12	13.86
	338469	3.09	5.11
10	338501	6.28	10.32
	338506	6.97	12.41
	338523	3.10	5.84
	338549	1.70	2.70
	338561	0.79	0.81
15	338662	1.72	1.46
	338671	0.17	0.91
	338676	2.10	15.86
	338726	1.20	1.09
	338779	0.12	0.57
20	338804	0.99	1.67
	338836	1.00	1.00
	338871	4.30	9.81
	338872	6.02	12.81
	338879	0.23	1.12
25	338937	6.55	12.26
	338966	1.76	5.42
	338993	1.00	2.40
	339047	5.26	10.81
	339100	5.10	6.88
30	339114	1.00	1.70
	339121	1.00	3.75
	339170	10.36	19.67
	339229	4.08	13.48
	339264	2.64	3.83
35	339293	1.73	1.94

TABLE 8B shows the accession numbers for those Pkeys in Table 8A lacking unigenes/D's. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
322044	187363_1	AW340926 AA249063 N86075
322060	44320_1	AI341937 AW003063 U34725 AA904742
321430	42705_1	X57414 X57415
321467	43034_1	X13075 X13076
322125	46779_1	R93901 AF075073 R93902
322166	46861_1	H69434 AF085958 H69846
322173	46873_1	H52567 H52557 AF085970 H52164
322178	46882_1	H56535 AF085980 H56712
322179	46885_1	H92891 AF085982 H92777
321577	1615102_1	H84849 H84252 H84260 H86684 H85320
321587	1615333_1	H95531 H95521 H84529
313723	111953_1	AA070412 AA102346 AA081885
320997	627492_1	H22544 H46842 AJ204929
322278	47271_1	W69304 AF086283 W69200
321687	218439_1	AA625149 AA313030 AA313052 H97463
313883	129439_1	AA665089 AA135130 AA484059 AA102419 AW877765
322320	47422_1	W79150 AF086419
322339	814584_1	AI668646 AI734214 W17348
314648	293660_1	AW979268 AA878419 AA431342 AA431628
300201	682222_1	AI308300 AI308296
306897	25196_2	AI093967
323155	979809_1	AL120701 AL135041 AL121524
322527	38927_1	AF147359 T58511 T58560
322585	473768_2	W88919 W89125
300362	1574395_1	Z42308 H23514
322635	82296_1	AA005129 AA679084 AA694399
322664	85042_1	AA011522 AA702841 AA011691 AA330797
315454	380580_1	AI239464 AI239473 AA625812 AI208703
322687	37372_1	AF074666 AI110759 AF090902
314852	327472_1	AI903735 AA491283 AI694953 AW976903 AA761362
307783	697809_1	AI347274 AW844024
324072	269032_1	AA381722 AA381829 AW963906 AW963902 AA381242
300627	221345_1	AA488472 W27363 AA317053 BE082689 AW967036 BE079872
323505	196389_1	AW970512 AA280261 AI652287 BE466438 AI650725 AA551854 AA281574 AW571481
315791	403558_1	AA678177 AA577034
324303	233842_1	AL118754 AA333202 H38001
316519	442885_1	AA847835 AA768378
300926	333127_1	AA504860 AA504911

	324580	328264_1	AA492588 AA492498 AA492571
	301882	275087_1	T78054 T79888 AA398185
	324804	398093_1	AI692552 AI393343 AI800510 AI377711 F24263 AA661876
5	324889	1515978_1	D31010 D30991 D31168 D31166 D31465
	302697	43219_1	AJ001409 AJ001410
	302711	45419_1	L08442 D51348
	302742	458_39	L12061
	318499	364430_1	T25451 AA585296 AA585305
10	310624	34624_4	U88896 U88898 AA916056 T03285 AI341594 AI359534 AI634031 U88897
	302847	458_105	X98941 X98942 X98943 X98953 X98949
	304122	77271_5	H28966
	303598	270283_1	AA382814 AA402411 AA412355
	311409	837264_1	AI698839 AI909260 AI909259
	312094	797889_1	Z78390 T97427
15	319312	1540116_1	Z45481 F12393 T74437
	319407	1688823_1	R05329 R01555 R08278
	319425	1689571_1	T82930 R02424 T85145
	320007	229683_1	AA336314 T82938 AA327744 AW967388 AA639967 T10753
20	320018	1815987_1	T83263 T85731 T85730
	319484	1691553_1	T91772 R07257 R07098
	318865	1535937_1	H10818 F07831 Z43072
	312220	1671607_1	N74613 T98756 T98589
	319546	243305_1	R09692 R09414 AA346353
	312389	902067_1	AI863140 W80703 R43474
25	319611	1566863_1	H14957 R56522 R11908
	312437	291472_1	BE080180 AW827313 AW231970 AA995028 AA428584 AW872716 AW892508 AW854593 AA578441 AW975234 AA664937 AA984131 AA528743 AA552874 AA564758 AW063245 AI267534 AW070190 AW893483 AA770330 AA906928 AA906582 AA758746 AA551717 AW063311 AA429538
30	311896	579192_1	AW205447 AI248530 AI084433 AI400976 R16553
	319834	112523_1	AA071267 T65940 T64515 AA071334
	321102	80531_1	AA018306 H38925 AA001221
	321158	410938_1	H79670 H47798 AA700289
	321199	212379_1	N34524 AA305071 AW954803 AA502335 AI433430 AI203597 AW026670 AW265323 AW850787 AA317554 AW993643 AW835572 AW385512 AI334966 W32951 H62656 H53902 R88904 AW835732
35	305528	28832_3	AA769156
	321270	1662057_1	N59537 N78278 R83560
	314126	177666_1	AA226431 AA226569 AA488748
	320714	743644_1	R91883 AI445591
40	306442	AA976899	
	306446	AA977348	
	306458	AA978186	
	306510	AA988546	
	306557	AA994530	
45	306572	AA995686	
	306582	AA996248	
	306656	AI004024	
	306686	AI015615	
	306751	AI032589	
50	308011	AI439473	
	306892	AI092465	
	308106	AI476803	
	308154	AI500500	
	308956	AI125111	
55	306958	AI125152	
	308213	AI557041	
	308216	AI557135	
	308219	AI557246	
	308588	AI718299	
	308599	AI719893	
60	308643	AI745040	
	308673	AI760864	
	308697	AI767143	
	308778	AI811109	
	308808	AI818289	
65	308875	AI832332	
	308886	AI833240	
	308898	AI858845	
	308966	AI870704	
	308979	AI873111	
70	303011	41689_1	AF090405 AF090407 AF090406
	303077	44060_1	AF163305 AF163307 AF163303
	305016	AA626876	
	305034	AA630128	
75	305072	AA641012	
	305148	AA654070	
	305190	AA665955	
	303978	AW513315	
	303990	AW515465	
	303998	AW516449	
80	303999	AW516611	
	305235	AA670480	
	305312	AA700201	
	305413	AA724659	
	305447	AA737856	
85	321244	29327_1	AF068654 AF068656 AF068655

	305614	AA782866	
	305637	AA806124	
	305639	AA806138	
5	305650	AA807709	
	305690	AA813477	
	305728	AA828209	
	305759	AA835353	
	305792	AA845256	
10	307041	AI144243	
	307091	AI167439	
	307181	AI189251	
	305901	AA872968	
	305910	AA875981	
15	307415	AI242118	
	307426	AI243364	
	307517	AI275055	
	307551	AI281556	
	307581	AI282207	
20	307608	AI290295	
	307691	AI318285	
	307730	AI336092	
	307760	AI342387	
	307764	AI342731	
25	307796	AI350556	
	309045	AI910902	
	309051	AI911975	
	307807	AI351799	
	307808	AI351826	
30	307820	AI355761	
	307852	AI365541	
	309122	AI928178	
	309164	AI937761	
	309177	AI951118	
35	307902	AI380462	
	309299	AW003478	
	309303	AW004823	
	309476	AW129368	
	309532	AW151119	
40	309747	AW264889	
	309769	AW272346	
	309799	AW276964	
	309866	AW299916	
	302679	311853_1	H55022 AA186889
45	309923	AW340684	
	309928	AW341418	
	309931	AW341683	
	309933	AW341936	
	302705	31765_1	U09060 U09061
50	302789	34161_1	AJ245067 AJ245070
	304006	AW517947	
	304024	T03036	
	304026	T03160	
	304028	T03266	
55	304046	T54803	
	304061	T61521	
	304063	T62536	
	302802	34487_1	Y08250 Y08245
	304114	R78946	
60	304155	H68696	
	304203	N56929	
	304234	W81608	
	304348	AA179868	
	304430	AA347682	
65	304456	AA411240	
	304521	AA464716	
	304526	AA476427	
	304607	AA513322	
	304735	AA576453	
70	304760	AA580401	
	305015	AA897116	
	306063	AA906316	
	306065	AA906725	
	306104	AA910956	
75	306109	AA911861	
	306242	AA932805	
	306288	AA936900	
	306396	AA970223	
	330568	NOT_FOUND_entrez	U56244
80	330599	15323_-12	U90437
	331131	genbank_R54797	R54797
	331203	NOT_FOUND_entrez	T82310
	331531	genbank_N51343	N51343
	331547	467396_1	AA828597 N54811
85	332074	genbank_AA599012	AA599012

TABLE 8C shows the genomic position for those Pkeys in Table 8A lacking unigene ID's and accession numbers. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

5	Pkey:	Unique number corresponding to an Eos probeset		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
10	Nt_position:	Indicates nucleotide positions of predicted exons.		
	Pkey	Ref	Strand	Nt_position
	332792	Dunham, I. et al.	Plus	73381-73768
	332816	Dunham, I. et al.	Plus	359844-360030
15	332906	Dunham, I. et al.	Plus	1923101-1923205
	332911	Dunham, I. et al.	Plus	1961767-1961858
	332912	Dunham, I. et al.	Plus	1962120-1962246
	332922	Dunham, I. et al.	Plus	2009620-2009738
	332956	Dunham, I. et al.	Plus	2510528-2510558
20	332959	Dunham, I. et al.	Plus	2518145-2518213
	333138	Dunham, I. et al.	Plus	3369205-3369323
	333139	Dunham, I. et al.	Plus	3369495-3369571
	333221	Dunham, I. et al.	Plus	3978070-3978187
	333380	Dunham, I. et al.	Plus	4904775-4904846
25	333387	Dunham, I. et al.	Plus	4910935-4910997
	333512	Dunham, I. et al.	Plus	5560510-5560564
	333524	Dunham, I. et al.	Plus	5612620-5612780
	333585	Dunham, I. et al.	Plus	6234778-6234894
	333618	Dunham, I. et al.	Plus	6562391-6562566
30	333627	Dunham, I. et al.	Plus	6620584-6620903
	333628	Dunham, I. et al.	Plus	6629004-6629233
	333650	Dunham, I. et al.	Plus	6796852-6797128
	333678	Dunham, I. et al.	Plus	7068223-7068288
	333750	Dunham, I. et al.	Plus	7608165-7608234
35	333763	Dunham, I. et al.	Plus	7692491-7692630
	333767	Dunham, I. et al.	Plus	7694407-7694623
	333768	Dunham, I. et al.	Plus	7695440-7695697
	333769	Dunham, I. et al.	Plus	7696625-7696707
	333772	Dunham, I. et al.	Plus	7706773-7706902
40	333777	Dunham, I. et al.	Plus	7746805-7746916
	333846	Dunham, I. et al.	Plus	8008623-8008757
	333884	Dunham, I. et al.	Plus	8153960-8154161
	333887	Dunham, I. et al.	Plus	8154882-8155025
	333891	Dunham, I. et al.	Plus	8156437-8156709
45	333892	Dunham, I. et al.	Plus	8156825-8157001
	333948	Dunham, I. et al.	Plus	8583497-8583627
	333954	Dunham, I. et al.	Plus	8563186-8563335
	333966	Dunham, I. et al.	Plus	8655643-8655826
	333968	Dunham, I. et al.	Plus	8681004-8681241
50	334061	Dunham, I. et al.	Plus	9686941-9687077
	334094	Dunham, I. et al.	Plus	9889953-9890105
	334113	Dunham, I. et al.	Plus	10282459-10282597
	334161	Dunham, I. et al.	Plus	10599033-10599180
55	334219	Dunham, I. et al.	Plus	12716160-12716384
	334239	Dunham, I. et al.	Plus	13056569-13056693
	334333	Dunham, I. et al.	Plus	13603544-13603657
	334378	Dunham, I. et al.	Plus	13907239-13907370
	334382	Dunham, I. et al.	Plus	13915866-13916036
60	334562	Dunham, I. et al.	Plus	14987847-14987940
	334588	Dunham, I. et al.	Plus	15032740-15032817
	334616	Dunham, I. et al.	Plus	15176123-15176470
	334633	Dunham, I. et al.	Plus	15333206-15333305
	334666	Dunham, I. et al.	Plus	18872214-18872317
	334891	Dunham, I. et al.	Plus	19299770-19299944
65	334934	Dunham, I. et al.	Plus	20103970-20104058
	335015	Dunham, I. et al.	Plus	20682792-20682945
	335120	Dunham, I. et al.	Plus	21436286-21436394
	335125	Dunham, I. et al.	Plus	21441390-21441471
	335179	Dunham, I. et al.	Plus	21634405-21634526
70	335188	Dunham, I. et al.	Plus	21669118-21669328
	335211	Dunham, I. et al.	Plus	21774611-21774680
	335361	Dunham, I. et al.	Plus	22807292-22807445
	335379	Dunham, I. et al.	Plus	22899306-22899420
	335414	Dunham, I. et al.	Plus	23235546-23235684
75	335416	Dunham, I. et al.	Plus	23237354-23237465
	335496	Dunham, I. et al.	Plus	24164386-24164545
	335497	Dunham, I. et al.	Plus	24167666-24167869
	335558	Dunham, I. et al.	Plus	24740167-24740347
	335586	Dunham, I. et al.	Plus	24990333-24990497
80	335686	Dunham, I. et al.	Plus	25439839-25439920
	335784	Dunham, I. et al.	Plus	25942710-25942792
	335823	Dunham, I. et al.	Plus	26365925-26366004
	335983	Dunham, I. et al.	Plus	27938968-27939070
	335995	Dunham, I. et al.	Plus	28009044-28009184
85	336021	Dunham, I. et al.	Plus	28686482-28686559

	336034	Dunham, I. et al.	Plus	29014404-29014590
	336038	Dunham, I. et al.	Plus	29022963-29023165
	336107	Dunham, I. et al.	Plus	299877731-29987869
5	336632	Dunham, I. et al.	Plus	983890-985529
	336633	Dunham, I. et al.	Plus	985591-985221
	336634	Dunham, I. et al.	Plus	986296-986670
	336635	Dunham, I. et al.	Plus	987908-988364
	336636	Dunham, I. et al.	Plus	988418-989185
10	336637	Dunham, I. et al.	Plus	989276-990813
	336638	Dunham, I. et al.	Plus	991906-993240
	336659	Dunham, I. et al.	Plus	1896402-1896478
	336694	Dunham, I. et al.	Plus	2420546-2420616
	336721	Dunham, I. et al.	Plus	3371522-3371586
15	336900	Dunham, I. et al.	Plus	10236423-10236523
	336948	Dunham, I. et al.	Plus	12692290-12692381
	337028	Dunham, I. et al.	Plus	16644817-16644942
	337054	Dunham, I. et al.	Plus	17821742-17821922
	337162	Dunham, I. et al.	Plus	23478943-23479145
20	337183	Dunham, I. et al.	Plus	23943606-23943696
	337184	Dunham, I. et al.	Plus	23973949-23974016
	337268	Dunham, I. et al.	Plus	28011979-28012034
	337299	Dunham, I. et al.	Plus	29022656-29022775
	337389	Dunham, I. et al.	Plus	31401509-31401579
25	337493	Dunham, I. et al.	Plus	33330760-33330981
	337549	Dunham, I. et al.	Plus	34474472-34474531
	337755	Dunham, I. et al.	Plus	3971764-3971900
	337809	Dunham, I. et al.	Plus	4449069-4449193
	337871	Dunham, I. et al.	Plus	5443027-5443101
30	337958	Dunham, I. et al.	Plus	6969162-6969270
	338008	Dunham, I. et al.	Plus	7697068-7697236
	338033	Dunham, I. et al.	Plus	8092128-8092271
	338110	Dunham, I. et al.	Plus	10384481-10384621
	338112	Dunham, I. et al.	Plus	10391398-10391600
35	338145	Dunham, I. et al.	Plus	11386529-11386692
	338148	Dunham, I. et al.	Plus	11448985-11449085
	338179	Dunham, I. et al.	Plus	12808775-12808833
	338197	Dunham, I. et al.	Plus	13638107-13638181
	338279	Dunham, I. et al.	Plus	16168944-16169091
40	338316	Dunham, I. et al.	Plus	17089711-17089988
	338322	Dunham, I. et al.	Plus	17132477-17132547
	338357	Dunham, I. et al.	Plus	18062184-18062402
	338359	Dunham, I. et al.	Plus	18074402-18074501
	338366	Dunham, I. et al.	Plus	18252026-18252189
45	338374	Dunham, I. et al.	Plus	18371200-18371282
	338414	Dunham, I. et al.	Plus	19345573-19345660
	338418	Dunham, I. et al.	Plus	19435506-19435596
	338501	Dunham, I. et al.	Plus	21244713-21244828
	338506	Dunham, I. et al.	Plus	21221871-21221953
50	338523	Dunham, I. et al.	Plus	21509763-21509864
	338662	Dunham, I. et al.	Plus	24404720-24404899
	338804	Dunham, I. et al.	Plus	27236005-27236108
	338836	Dunham, I. et al.	Plus	27792166-27792272
	338879	Dunham, I. et al.	Plus	28410653-28410734
55	338937	Dunham, I. et al.	Plus	29160655-29160725
	338953	Dunham, I. et al.	Plus	30077787-30078184
	339047	Dunham, I. et al.	Plus	30760793-30760968
	339100	Dunham, I. et al.	Plus	31141580-31141765
	339114	Dunham, I. et al.	Plus	31456454-31456519
60	339121	Dunham, I. et al.	Plus	31583467-31583536
	339170	Dunham, I. et al.	Plus	32216399-32216527
	339293	Dunham, I. et al.	Plus	33223671-33223819
	332858	Dunham, I. et al.	Minus	1339607-1339397
	332982	Dunham, I. et al.	Minus	2628286-2628109
65	332984	Dunham, I. et al.	Minus	2632606-2632457
	332998	Dunham, I. et al.	Minus	2711704-2711565
	333058	Dunham, I. et al.	Minus	3028925-3028811
	333097	Dunham, I. et al.	Minus	3204124-3204036
	333121	Dunham, I. et al.	Minus	3308446-3308358
	333122	Dunham, I. et al.	Minus	3309596-3309531
70	333123	Dunham, I. et al.	Minus	3310817-3310749
	333140	Dunham, I. et al.	Minus	3377220-3376309
	333260	Dunham, I. et al.	Minus	4308400-4308304
	333603	Dunham, I. et al.	Minus	6465335-6465727
75	333604	Dunham, I. et al.	Minus	6467090-6466768
	333904	Dunham, I. et al.	Minus	8217374-8217261
	333906	Dunham, I. et al.	Minus	8218238-8218063
	334183	Dunham, I. et al.	Minus	11832582-11832508
	334187	Dunham, I. et al.	Minus	11921456-11921205
80	334222	Dunham, I. et al.	Minus	12732417-12732289
	334223	Dunham, I. et al.	Minus	12734365-12734269
	334255	Dunham, I. et al.	Minus	13200776-13200692
	334492	Dunham, I. et al.	Minus	14478333-14478172
	334648	Dunham, I. et al.	Minus	15363301-15363222
85	334787	Dunham, I. et al.	Minus	16299093-16298937
	334933	Dunham, I. et al.	Minus	20078117-20077991

WO 02/086443			
	334945	Dunham, I. et al.	Minus 20138885-20138837
	334957	Dunham, I. et al.	Minus 20173311-20173218
	334990	Dunham, I. et al.	Minus 20341159-20341087
5	335093	Dunham, I. et al.	Minus 21297367-21297214
	335288	Dunham, I. et al.	Minus 22304275-22303770
	335289	Dunham, I. et al.	Minus 22305950-22305708
	335548	Dunham, I. et al.	Minus 24662773-24662673
	335551	Dunham, I. et al.	Minus 24679828-24678961
10	335619	Dunham, I. et al.	Minus 25082677-25082498
	335620	Dunham, I. et al.	Minus 25092561-25092434
	335621	Dunham, I. et al.	Minus 25098878-25098767
	335682	Dunham, I. et al.	Minus 25421215-25421093
	335755	Dunham, I. et al.	Minus 25763806-25763747
15	335814	Dunham, I. et al.	Minus 26320043-26319845
	335815	Dunham, I. et al.	Minus 26320518-26320421
	335835	Dunham, I. et al.	Minus 26393311-26393245
	335851	Dunham, I. et al.	Minus 26604863-26604742
	335868	Dunham, I. et al.	Minus 26711437-26711300
20	335896	Dunham, I. et al.	Minus 26977639-26977558
	335936	Dunham, I. et al.	Minus 27360474-27360400
	335948	Dunham, I. et al.	Minus 27555924-27555788
	336066	Dunham, I. et al.	Minus 29241080-29240842
	336205	Dunham, I. et al.	Minus 30477456-30477311
25	336275	Dunham, I. et al.	Minus 32086675-32086536
	336292	Dunham, I. et al.	Minus 32818035-32817927
	336331	Dunham, I. et al.	Minus 33594527-33594371
	336419	Dunham, I. et al.	Minus 34052568-34052445
	336675	Dunham, I. et al.	Minus 2020758-2020664
30	336684	Dunham, I. et al.	Minus 2158060-2157993
	336716	Dunham, I. et al.	Minus 3259952-3259862
	336798	Dunham, I. et al.	Minus 5888954-5888757
	337043	Dunham, I. et al.	Minus 17407330-17407251
	337046	Dunham, I. et al.	Minus 17610892-17610821
35	337128	Dunham, I. et al.	Minus 22215251-22215034
	337192	Dunham, I. et al.	Minus 24591853-24591771
	337194	Dunham, I. et al.	Minus 24610510-24610359
	337229	Dunham, I. et al.	Minus 26716579-26716481
	337325	Dunham, I. et al.	Minus 30015948-30015800
40	337497	Dunham, I. et al.	Minus 33371317-33371258
	337500	Dunham, I. et al.	Minus 33376212-33376158
	337603	Dunham, I. et al.	Minus 1299296-1299194
	337605	Dunham, I. et al.	Minus 1346555-1346397
	337671	Dunham, I. et al.	Minus 3260634-3260547
45	337786	Dunham, I. et al.	Minus 4133203-4133081
	337862	Dunham, I. et al.	Minus 5347658-5347550
	338083	Dunham, I. et al.	Minus 9318438-9318301
	338158	Dunham, I. et al.	Minus 11794465-11794343
	338161	Dunham, I. et al.	Minus 12124716-12124658
50	338182	Dunham, I. et al.	Minus 12824919-12824827
	338189	Dunham, I. et al.	Minus 12878594-12878478
	338199	Dunham, I. et al.	Minus 13760865-13760780
	338215	Dunham, I. et al.	Minus 14055447-14055355
	338469	Dunham, I. et al.	Minus 20520387-20520242
55	338549	Dunham, I. et al.	Minus 22049171-22049081
	338561	Dunham, I. et al.	Minus 22311966-22311856
	338671	Dunham, I. et al.	Minus 24508421-24508346
	338676	Dunham, I. et al.	Minus 24637427-24637369
	338726	Dunham, I. et al.	Minus 25926206-25925618
50	338779	Dunham, I. et al.	Minus 27030151-27029795
	338871	Dunham, I. et al.	Minus 28301708-28301611
	338872	Dunham, I. et al.	Minus 28300921-28300790
	338966	Dunham, I. et al.	Minus 29614876-29614749
	339229	Dunham, I. et al.	Minus 32722330-32722199
55	339264	Dunham, I. et al.	Minus 32975145-32975053
	325228	6381940 Plus	2630-2694
	325235	6381943 Minus	162154-162264
	329588	3952484 Plus	1169-1819
	329560	3952491 Plus	2095-2990
70	329541	3983503 Minus	2765-3059
	325328	5866875 Plus	86780-86854
	325340	6017033 Minus	166656-166819
	325373	5866920 Minus	1136686-1136777
	325367	5866920 Minus	922881-922958
	325389	5866921 Plus	239672-239759
75	325436	5866939 Minus	29778-29907
	325498	5866967 Plus	173372-173930
	325471	6017034 Minus	289268-289342
	325557	6056302 Plus	50921-51050
30	325559	6249595 Minus	118590-119172
	325560	6249595 Minus	133794-133981
	325569	6249599 Plus	79927-80217
	325587	6682462 Plus	126724-126967
	325585	6682462 Plus	73476-73574
	325597	5866992 Plus	1065020-1065089
35	325639	5867002 Plus	253525-253608

	325739	5867038	Minus	205139-205269
	325740	5867038	Minus	207533-207690
	325792	6469828	Minus	1018-1176
5	325735	6552447	Minus	269122-269190
	325685	6682468	Plus	117397-117483
	325686	6682468	Plus	118337-118439
	325819	6682490	Minus	130314-130370
	329764	6048195	Minus	109733-109968
10	329703	6065793	Minus	139994-140138
	326543	6448539	Plus	53403-53537
	329816	6624888	Minus	70296-70423
	329860	6687260	Minus	163474-163605
	325883	5867087	Plus	22498-22663
	325895	5867097	Plus	358317-358476
15	325925	5867124	Plus	115749-115962
	325932	5867127	Plus	7369-7441
	325941	5867133	Minus	64228-64402
	325969	5867153	Plus	101911-102081
	325971	5867153	Plus	105841-106035
20	329993	4567166	Minus	101307-101434
	330020	6671887	Plus	172397-172491
	326163	5867168	Minus	7831-8035
	326274	5867171	Minus	410289-410404
	326025	5867176	Plus	70854-70915
25	326046	5867182	Minus	62668-62825
	326099	5867186	Minus	661381-661510
	326108	5867187	Minus	23784-23903
	326165	5867208	Minus	62787-62929
30	326189	5867212	Plus	69288-69413
	326204	5867218	Minus	148088-148200
	326230	5867230	Minus	301868-301972
	330052	4567182	Plus	352560-352963
	330036	6042048	Plus	117120-117216
35	326360	5867293	Plus	13627-13844
	326589	5867320	Plus	22760-22919
	326393	5867341	Plus	41702-41841
	326505	5867435	Minus	8818-8949
	326515	5867439	Plus	36683-36809
40	326592	6138928	Plus	23689-23828
	330107	6015249	Minus	100091-100282
	330106	6015249	Minus	99443-99778
	330100	6015253	Plus	21166-21301
	330093	6015278	Plus	1043-1199
45	330088	6015293	Plus	37517-37638
	330085	6015302	Minus	59613-59770
	330120	6671864	Minus	127553-127656
	330123	6671869	Minus	35311-35406
	326742	5867611	Minus	95187-95248
50	326805	5867637	Plus	24656-24749
	326818	6117831	Minus	15199-15309
	326720	6552456	Plus	84525-84677
	326770	6598307	Minus	513603-513668
	326692	6682502	Plus	117697-117899
55	326693	6682502	Minus	335002-335095
	326983	5867657	Minus	16023-16681
	326991	5867660	Plus	18147-18339
	326936	6004446	Minus	10217-10357
	326964	6469836	Plus	75340-75456
60	327040	6531965	Plus	783670-783817
	327053	6531965	Plus	2247267-2247437
	327075	6531965	Plus	4041318-4041431
	327085	6531965	Plus	4734947-4735069
	327036	6531965	Plus	319951-320040
65	327130	6531976	Plus	20247-22343
	327155	5866841	Minus	2462-2620
	327288	5867481	Plus	48583-48773
	327332	5867516	Minus	56361-56532
	327220	5867525	Minus	65701-65781
70	327224	5867534	Plus	188468-188544
	327321	6249562	Minus	99745-99836
	327361	6552412	Minus	61013-62130
	327396	5867743	Plus	8702-8820
	327414	5867750	Plus	102461-102586
75	327442	5867759	Plus	111483-111618
	327457	5867772	Plus	88030-88151
	327473	5867775	Plus	75101-75181
	327483	5867783	Plus	181573-181662
	327377	5867793	Minus	37610-37676
80	327562	5867804	Minus	343989-344474
	327568	5867811	Minus	46152-46287
	327606	6004463	Plus	200262-200495
	327611	5867868	Minus	175063-175392
	327642	5867891	Minus	2513-2743
85	327654	5867910	Minus	97564-97710
	327734	5867940	Minus	31003-31583

	327775	5867964	Minus	130791-130871
	327796	5867982	Plus	85267-85405
	327840	6249578	Minus	73065-73206
5	330208	6013599	Plus	66517-66931
	330263	6671884	Minus	101503-101634
	328004	5867993	Minus	157407-157887
	328101	5868020	Plus	289920-290014
	328100	5868020	Minus	263545-263635
10	328113	5868024	Minus	80378-80491
	328157	5868064	Plus	73326-73615
	328196	5868080	Minus	16551-16729
	328197	5868081	Minus	42133-42438
	327940	5868197	Minus	95240-95428
15	327984	5868216	Plus	66611-66677
	328021	5902482	Plus	713478-714590
	328068	6117819	Plus	253903-254022
	328264	6381912	Plus	55086-55404
	330300	2905862	Minus	3246-3302
20	328608	5868222	Minus	87770-87953
	328600	5868229	Minus	38889-40010
	328616	5868239	Plus	293920-294224
	328623	5868246	Minus	120020-120126
	328632	5868247	Plus	76734-76853
25	328666	5868254	Minus	778-901
	328698	5868264	Minus	625555-625633
	328700	5868264	Plus	764089-764203
	328708	5868271	Minus	68114-68854
	328735	5868289	Plus	89389-89455
30	328743	5868289	Plus	274638-274726
	328806	5868324	Plus	29408-29684
	328299	5868366	Minus	149708-149889
	328342	5868363	Plus	59955-60094
	328365	5868387	Minus	270724-270798
35	328369	5868388	Plus	75371-75583
	328381	5868392	Plus	662758-662848
	328451	5868425	Minus	217275-217336
	328481	5868449	Minus	8987-9180
	328500	5868464	Plus	59098-59481
40	328530	5868482	Plus	334973-335408
	328654	6004473	Plus	1193739-1193866
	328861	6381928	Minus	108317-108403
	328908	5868493	Plus	117002-117059
	328933	5868500	Plus	771755-771889
45	328934	5868500	Plus	846342-846448
	328949	6456765	Minus	43552-43619
	330313	6042030	Minus	33642-33775
	329005	5868542	Plus	85470-85673
	330366	2944106	Plus	151837-151914
50	330372	6580495	Minus	317461-317688
	329033	5868561	Minus	5390-5479
	329037	5868562	Minus	32466-32562
	329067	5868591	Minus	146417-147652
	329134	5868679	Plus	29959-30018
55	329157	5868687	Minus	145940-146155
	329178	5868704	Plus	179177-179463
	329192	5868716	Plus	166936-167020
	329194	5868716	Minus	304450-304559
	329204	5868720	Minus	3050-3190
60	329224	5868728	Plus	27422-27664
	329228	5868728	Minus	50118-50287
	329288	5868771	Plus	25554-26299
	329337	5868806	Minus	467155-467222
	329011	6682532	Plus	48658-48741

TABLE 9A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer

Table 9A shows about 1312 genes up-regulated in lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) relative to normal body tissues. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 9B show the accession numbers for those Pkey's lacking UnigenelD's for table 9A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 9C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 9A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset Identifier number
ExAcct: Exemplar Accession number, Genbank accession number
UnigenelD: UnigenelD number
Unigenel Title: UnigenelD gene title
R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAcct	UnigenelD	Unigenel Title	R1	R2
400195			NM_007057:Homo sapiens ZW10 interactor	1.00	1.00
400205			NM_006265:Homo sapiens RAD21 (S. pombe)	15.80	396.00
400220			Eos Control	2.28	2.84
400277			Eos Control	7.68	9.72
400285			Eos Control	1.00	1.00
400288	X06256	Hs.149609	integrin, alpha 5 (fibronectin receptor,	1.04	2.24
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.85	74.00
400301	X03635	Hs.1657	estrogen receptor 1	1.00	1.00
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	1.75	1.65
400328	X87344	Hs.180062	transporter 2, ATP-binding cassette, sub	0.87	1.80
400419	AF084545		Target	156.55	253.00
400512			NM_030878:Homo sapiens cytochrome P450,	1.00	2.00
400517	AF242388		lengsin	3.67	87.00
400560			NM_030878:Homo sapiens cytochrome P450,	1.00	1.00
400664			NM_002425:Homo sapiens matrix metallopro	20.26	45.00
400665			NM_002425:Homo sapiens matrix metallopro	1.36	1.07
400666			NM_002425:Homo sapiens matrix metallopro	3.26	3.22
400749			NM_003105:Homo sapiens sortilin-related	1.00	91.00
400763			Target Exon	7.63	24.00
401027			Target Exon	1.00	1.00
401093			C12000586:gij6330167 dbj BAA86477.1 (A	1.00	155.00
401203			Target Exon	1.00	86.00
401212			C12000457:gij7512178 pit T30337 polypr	1.00	400.00
401411			ENSP00000247172:HYPOTHETICAL 126.2 kDa	1.00	72.00
401435			C14000397:gij749898 pit T33295 hypoth	1.00	64.00
401464	AF039241		histone deacetylase 5	3.82	49.00
401714			ENSP00000241802:CDNA FLJ11007 FIS, CLON	2.02	40.00
401747			Homo sapiens keratin 17 (KRT17)	128.43	68.00
401760			Target Exon	1.74	35.00
401780			NM_005557:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275:Homo sapiens keratin 15 (KRT1	4.13	2.70
401797			Target Exon	1.44	2.10
401961			NM_021626:Homo sapiens serine carboxypep	1.41	1.86
401985	AF053004		class I cytokine receptor	1.00	177.00
401994			Target Exon	61.84	47.00
402075			ENSP00000251056:Plasma membrane calcium	1.00	1.00
402260			NM_001436:Homo sapiens fibrillar (FBL	1.58	1.39
402265			Target Exon	2.09	35.00
402297			Target Exon	1.00	92.00
402408			NM_030920:Homo sapiens hypothetical pro	28.87	13.00
402420			C1000823:gij1043240 emb CAC10290.1 (A	1.00	1.44
402674			Target Exon	7.44	243.00
402802			NM_001397:Homo sapiens endothelin conver	1.00	70.00
402994			NM_002463:Homo sapiens myxovirus (influ	1.37	1.43
403137			NM_005381:Homo sapiens nucleolin (NCL),	1.00	19.00
403306	NM_006825		transmembrane protein (53kD), endoplasmic	1.00	43.00
403329			Target Exon	1.00	61.00
403381			ENSP00000231844:Ecotropic virus integra	1.00	119.00
403478			NM_022342:Homo sapiens kinesin protein 9	28.13	136.00
403485			C3001813:gij12737279 ref XP_012163.1 k	20.23	76.00
403627			Target Exon	6.30	29.33
403715			Target Exon	1.30	35.00
404044			ENSP00000237855:DJ398G3.2 (NOVEL PROTEI	1.00	54.00
404076			NM_016020:Homo sapiens CGI-75 protein (14.29	91.00
404101			C8000950:gij423560 pit A47318 RNA-bindi	1.00	1.00
404140			NM_006510:Homo sapiens ret finger protei	1.42	1.44
404165			ENSP00000244562:NRH dehydrogenase [quinol	1.00	54.00
404185			Target Exon	1.00	117.00
404210			NM_005936:Homo sapiens myeloid/lymphoid	5.93	13.77
404253			NM_021058:Homo sapiens H2B histone fami	1.00	1.00

404287		C6001909:gil704441[dbj]BAA18909.1] (D298	29.71	42.00
404298		C6001236*:gil121715[sp]P26697[MTA3_CHICK	1.30	1.00
404347		Target Exon	1.00	1.00
404440		NM_021048:Homo sapiens melanoma antigen,	1.00	15.00
404721		NM_005596*:Homo sapiens nuclear factor I	1.00	60.00
404794	NM_000078	cholesteryl ester transfer protein, plas	1.07	1.38
404854		Target Exon	1.61	2.01
404877		NM_005365:Homo sapiens melanoma antigen,	1.00	1.00
404927		Target Exon	1.00	1.00
404996		Target Exon	1.00	1.00
405449		CY000047*:gil11427234[ref]XP_009399.1] z	1.00	1.00
405568		NM_031413*:Homo sapiens cat eye syndrome	1.00	78.00
405572		Target Exon	0.76	1.14
405646		C1200200:gil4557225[ref]NP_000005.1] al	1.01	1.28
405676	BE336714	cytochrome c-1	1.13	2.89
405770		NM_002362:Homo sapiens melanoma antigen,	45.52	37.00
405932		C15000305:gil3806122[gb]AAC69198.1] (AFO	1.99	1.99
406137		NM_000179*:Homo sapiens mutS (E. coli) h	2.77	2.38
406350		Target Exon	1.00	35.00
406399		NM_003122*:Homo sapiens serine protease	1.00	39.00
406467		Target Exon	1.00	1.00
406621	X57809	Immunoglobulin lambda locus	1.41	1.74
406642	AJ245210	gb:Homo sapiens mRNA for immunoglobulin	2.16	3.91
406663	U24683	Immunoglobulin heavy constant mu	2.07	2.93
406671	AA129547	met proto-oncogene (hepatocyte growth fa	15.00	51.00
406673	M34996	major histocompatibility complex, class	0.98	3.09
406676	X58399	Human L2.9 transcript of unrearranged im	1.30	1.53
406678	U77534	gb:Human clone 1A11 immunoglobulin varia	1.33	1.45
406685	M18728	gb:Human nonspecific crossreacting antig	1.46	2.85
406687	M31126	pregnancy specific beta-1-glycoprotein 9	8.61	8.50
406690	M29540	carcinoembryonic antigen-related cell ad	226.37	350.00
406698	X03058	major histocompatibility complex, class	1.01	2.52
406815	AA833930	tRNA isopentenylpyrophosphate transferas	20.25	32.00
406851	AA609784	major histocompatibility complex, class	0.75	1.91
406964	M21305	gb:Human alpha satellite and satellite 3	38.15	1114.00
406967	M24349	gb:Human parathyroid hormone-like protei	1.00	1.00
406974	M57293	gb:Human parathyroid hormone-related pep	1.00	1.00
407103	AA424881	hypothetical protein MGC13170	1.77	1.10
407128	R83312	EST	1.00	1.00
407137	T97307	gb:ye53h05.s1 Soares fetal liver spleen	142.70	135.00
407168	R45175	ESTs	2.16	18.00
407239	AA076350	leukocyte immunoglobulin-like receptor,	1.10	1.57
407242	M18728	gb:Human nonspecific crossreacting antig	1.12	2.85
407244	M10014	fibrinogen, gamma polypeptide	3.24	15.38
407289	AA135159	Homo sapiens cDNA FLJ12149 fis, clone MA	3.53	3.68
407300	AA102616	gb:zn43e07.s1 Stratagene HeLa cell s3 93	19.74	73.00
407366	AF026942	gb:Homo sapiens cig33 mRNA, partial sequ	0.06	8.25
407378	AA299264	ESTs, Moderately similar to I38022 hypot	1.00	26.00
407430	AF169351	gb:Homo sapiens protein tyrosine phospho	1.00	25.00
407453	AJ132087	gb:Homo sapiens mRNA for axonemal dynein	1.00	75.00
407577	AW131324	hypothetical protein MGC12538	1.00	1.00
407634	AW016569	UDP-GlcNAc:betaGal beta-1,3-N-acetylgluc	111.20	228.00
407710	AW022727	ESTs	1.00	28.00
407720	AB037776	KIAA1355 protein	1.89	1.31
407746	AK001962	hypothetical protein FLJ11100	1.00	1.00
407756	AA116021	ubiquitin specific protease 18	4.51	5.00
407758	D50915	KIAA0125 gene product	1.00	28.00
407782	AA608956	ESTs, Moderately similar to PURKINJE CEL	0.97	1.14
407788	BE514982	S100 calcium-binding protein A2	7.88	3.83
407790	AI027274	Homo sapiens cDNA FLJ14866 fis, clone PL	3.63	42.00
407811	AW190902	cysteine knot superfamily 1, BMP antagon	89.96	109.00
407839	AA045144	ESTs	173.91	108.00
407944	R34008	desmocollin 2	111.30	70.00
408000	L11690	bullous pemphigoid antigen 1 (230/240kD)	151.17	8.00
408031	AA081395	Homo sapiens cDNA FLJ10366 fis, clone NT	9.91	93.00
408053	BE086548	calcineurin-binding protein calcyphorin-1	195.78	231.00
408070	AW148852	gb:xf05d05.x1 NCI_CGAP_Bm35 Homo sapien	1.00	1.00
408101	AW968504	CDC2-related protein kinase 7	37.84	61.00
408122	AI432652	hypothetical protein FLJ10718	0.85	1.71
408212	AA297567	hypothetical protein	5.88	7.91
408243	Y00787	interleukin 8	4.27	9.98
408349	BE546947	homeo box C10	3.79	3.46
408353	BE439838	mitochondrial ribosomal protein S17	1.88	1.65
408354	AJ382803	ESTs	1.00	73.00
408369	R38438	solute carrier family 15 (H777) transport	1.41	16.50
408380	AF123050	diubiquitin	15.19	37.22
408482	NM_000676	adenosine A2b receptor	1.65	1.19
408522	AI541214	Small proline-rich protein SPRK [human,	1.98	1.24
408536	AW381532	ESTs	1.55	1.50
408545	AW235405	ESTs	1.00	1.00
408572	AA055611	ESTs, Moderately similar to ALU4_HUMAN A	1.00	44.00
408633	AW963372	PRO2000 protein	107.16	58.00
408660	AA525775	ESTs, Moderately similar to PC4259 ferri	1.00	1.00
408761	AA057264	ESTs, Weakly similar to (define not ava	52.24	141.00
408771	AW732573	potassium voltage-gated channel, delayed	3.05	109.00

	408783	AF192522	Hs.47701	NPC1 (Niemann-Pick disease, type C1, gen	1.02	1.07
	408790	AW580227	Hs.47850	neurotrophic tyrosine kinase, receptor,	41.19	61.00
	408805	H59912	Hs.48269	vaccinia related kinase 1	24.67	45.00
5	408841	AW438865	Hs.256862	ESTs	1.00	58.00
	408873	AL046017	Hs.182278	calmodulin 2 (phosphorylase kinase, del	1.00	89.00
	408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
	408992	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.00	1.00
	408996	AI979168	Hs.344096	glycoprotein (transmembrane) nmb	3.71	5.50
10	409015	BE389387	Hs.49767	NM_004553:Homo sapiens NADH dehydrogenas	1.44	1.24
	409038	T97490	Hs.50002	small inducible cytokine subfamily A (Cy	4.28	5.32
	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KAA119	112.42	195.00
	409077	AA401369	Hs.190721	ESTs	1.00	17.00
	409093	BE243834	Hs.50441	CGI-04 protein	2.02	1.93
15	409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
	409142	AL136877	Hs.50758	SMC4 (structural maintenance of chromoso	14.87	6.00
	409187	AF154830	Hs.50966	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
	409228	AI654298	Hs.271695	ESTs, Weakly similar to 2109260A B cell	1.22	1.00
	409234	AI879419	Hs.27206	ESTs	1.00	1.00
20	409268	AA625304	Hs.187579	ESTs	11.90	23.00
	409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	168.91	35.00
	409404	BE220053	Hs.129056	ESTs	1.00	1.00
	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	79.74	96.00
25	409430	R21945	Hs.346735	splicing factor, arginine/serine-rich 5	1.45	2.10
	409446	AI581173	Hs.67688	ESTs	1.00	4.00
	409506	NM_006153	Hs.54589	NCK adaptor protein 1	3.97	28.00
	409522	AA075382		gb:zm87b03.s1 Stratagene ovarian cancer	15.98	141.00
	409582	AA401369	Hs.190721	ESTs	1.00	17.00
30	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	292.12	79.00
	409705	M37762	Hs.56023	brain-derived neurotrophic factor	1.00	82.00
	409719	AI769160	Hs.108681	Homo sapiens brain tumor associated prot	1.00	1.00
	409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	0.12	18.12
	409744	AW675258	Hs.56265	Homo sapiens mRNA; cDNA DKFZp586P2321 (f	20.75	51.00
35	409757	NM_001899	Hs.123114	cystatin SN	22.46	15.80
	409866	AW502152		gb:U1-HF-BR0p-ajr-f-11-0-U1.r1 NIH_MGC_5	1.00	1.00
	409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	1.50	1.09
	409902	AI337658	Hs.156351	ESTs	25.92	50.00
	409935	AW511413	Hs.278025	ESTs	2.63	2.11
40	409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	2.17	4.01
	409958	NM_001523	Hs.57697	hyaluronan synthase 1	0.91	2.07
	410001	AB041036	Hs.57771	kallikrein 11	1.04	2.28
	410032	BE065985		gb:RC3-BT0319-120200-014-a09 BT0319 Homo	1.00	58.00
	410037	AB020725	Hs.58009	KIAA0918 protein	1.00	34.00
45	410044	BE566742	Hs.58169	highly expressed in cancer, rich in leuc	1.00	1.00
	410048	W76467	Hs.58218	proline oxidase homolog	1.03	1.44
	410076	T05387	Hs.7991	ESTs	1.12	1.50
	410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
	410153	BE311926	Hs.15830	hypothetical protein FLJ12691	1.00	1.00
50	410166	AK001376	Hs.59346	hypothetical protein FLJ10514	1.00	1.00
	410193	AJ132592	Hs.59757	zinc finger protein 281	42.01	51.00
	410274	AA381807	Hs.61762	hypoxia-inducible protein 2	1.72	1.32
	410309	BE043077	Hs.278153	ESTs	1.00	2.00
	410340	AW182833	Hs.112188	hypothetical protein FLJ13149	32.08	75.00
55	410348	AW182663	Hs.95469	ESTs	1.00	1.00
	410407	X66839	Hs.63287	carbonic anhydrase IX	1.40	1.11
	410418	D31382	Hs.63325	transmembrane protease, serine 4	4.30	2.03
	410438	AB037756	Hs.45207	hypothetical protein KIAA1335	1.00	18.00
	410553	AW016824	Hs.255527	hypothetical protein MGC14128	1.34	1.04
60	410555	W27235	Hs.64311	a disintegrin and metalloproteinase doma	23.99	1.41
	410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
	410681	AW246890	Hs.65425	calbindin 1, (28kD)	10.88	18.92
	410781	AI375672	Hs.165028	ESTs	1.00	57.00
	411027	AF072099	Hs.67846	leukocyte immunoglobulin-like receptor,	1.62	3.78
65	411074	X60435	Hs.68137	adenylate cyclase activating polypeptide	1.00	1.15
	411089	AA456454		cell division cycle 2-like 1 (PITSLRE pr	1.56	1.58
	411152	BE069199		gb:QV3-BT0379-010300-105-g03 BT0379 Homo	1.00	84.00
	411248	AA551538	Hs.334605	Homo sapiens cDNA FLJ14409 fis, clone HE	1.82	1.45
	411252	AB018549	Hs.69328	MD-2 protein	7.32	12.74
70	411263	BE297802	Hs.69360	kinesin-like 6 (mitotic centromere-assoc	3.44	2.55
	411365	M76477	Hs.289082	GM2 ganglioside activator protein	1.35	2.02
	411402	BE297855	Hs.69855	NRAS-related gene	1.00	46.00
	411573	AB029000	Hs.70823	KIAA1077 protein	11.40	11.35
	411579	AC005258	Hs.70830	U8 snRNA-associated Sm-like protein LSm7	1.08	1.90
75	411617	AA247994	Hs.90063	neurocalcin delta	1.74	2.57
	411732	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.02	1.00
	411773	NM_006799	Hs.72026	protease, serine, 21 (testisin)	1.34	2.19
	411789	AF245505	Hs.72157	Adican	2.19	2.79
	411800	N39342	Hs.103042	microtubule-associated protein 18	23.34	34.00
80	411945	AL033527	Hs.92137	v-myc avian myelocytomatosis viral oncog	1.00	8.00
	412115	AK001763	Hs.73239	hypothetical protein FLJ10901	2.07	1.64
	412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rabkinas	118.48	92.00
	412276	BE262621	Hs.73798	macrophage migration inhibitory factor (1.98	1.49
	412464	T78141	Hs.22826	ESTs, Weakly similar to I55214 salivary	1.16	1.34
85	412530	AA766268	Hs.266273	hypothetical protein FLJ13346	41.52	84.00
	412537	AL031778		nuclear transcription factor Y, alpha	17.90	55.00

	412659	AW753865	Hs.74376	olfactomedin related ER localized protei	14.65	47.00
	412719	AW016610	Hs.816	ESTs	382.46	128.00
	412723	AA648459	Hs.335951	hypothetical protein AF301222	54.90	1.00
	412811	H06382		ESTs	1.00	11.00
5	412817	AL037159	Hs.74619	proteasome (prosome, macropain) 26S subu	1.63	1.42
	412863	AA121673	Hs.59757	zinc finger protein 281	17.63	56.00
	412924	BE018422	Hs.75258	H2A histone family, member Y	1.00	22.00
	413004	T35901	Hs.75117	interleukin enhancer binding factor 2, 4	2.19	2.05
10	413011	AW068115	Hs.821	biglycan	1.22	1.88
	413048	M93221	Hs.75182	mannose receptor, C type 1	0.30	6.23
	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote	3.43	8.71
	413129	AF292100	Hs.104613	RP42 homolog	4.67	4.77
	413142	M81740	Hs.75212	ornithine decarboxylase 1	1.92	2.59
	413223	A1732182	Hs.191866	ESTs	5.73	27.00
15	413248	T64858	Hs.21433	hypothetical protein DKFZp547J036	0.99	1.06
	413273	U75679	Hs.75257	stem-loop (histone) binding protein	1.00	18.00
	413278	BE563085	Hs.833	interferon-stimulated protein, 15 kDa	1.10	1.09
	413281	AA861271	Hs.222024	transcription factor BMAL2	95.94	69.00
20	413364	BE536218	Hs.137516	fidgetin-like 1	1.00	1.00
	413385	M34455	Hs.840	indoleamine-pyrrole 2,3 dioxygenase	0.95	2.09
	413409	AJ638418	Hs.1440	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	1.00	1.00
	413453	AA129640	Hs.128065	ESTs	1.00	31.00
	413527	BE250788	Hs.179882	hypothetical protein FLJ12443	1.08	1.46
	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	79.15	114.00
25	413573	AJ733859	Hs.149089	ESTs	1.00	1.00
	413582	AW295647	Hs.71331	hypothetical protein MGC5350	8.80	10.00
	413597	AW302885	Hs.117163	ESTs	1.00	1.00
	413690	BE157489		gb:RC1-HT0375-120200-011-e06 HT0375 Homo	1.00	1.00
30	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B	3.16	2.32
	413719	BE439580	Hs.75498	small inducible cytokine subfamily A (Cy	2.88	9.52
	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kafirin	144.10	108.00
	413801	M62246	Hs.35406	ESTs, Highly similar to unnamed protein	1.00	17.00
	413833	Z15005	Hs.75573	centromere protein E (312kD)	1.00	1.00
35	413882	AA132973	Hs.184492	ESTs	64.24	148.00
	413926	AA133338	Hs.54310	ESTs	1.00	67.00
	413943	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	43.42	42.00
	413995	BE048146	Hs.75571	synaptobrevin 1A (brain)	1.23	1.11
	414035	Y00630	Hs.75716	serine (or cysteine) proteinase inhibito	2.02	2.51
40	414142	AW368397	Hs.334485	Homo sapiens cDNA FLJ114438 fis, clone HE	1.00	102.00
	414180	AJ863304	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	6.92	77.00
	414245	BE148072	Hs.75850	WAS protein family, member 1	1.00	1.00
	414275	AW970254	Hs.889	Charot-Leyden crystal protein	1.00	59.00
	414317	BE263280	Hs.75888	phosphogluconate dehydrogenase	1.52	1.73
45	414334	AA824298	Hs.21331	hypothetical protein FLJ10036	1.78	1.72
	414341	D80004	Hs.75909	KIAA0182 protein	33.90	151.00
	414368	W70171	Hs.75939	uridine monophosphate kinase	171.60	97.00
	414416	AW409985	Hs.76084	hypothetical protein MGC2721	2.32	1.85
	414430	AJ346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1	226.15	66.00
50	414570	Y00285	Hs.76473	insulin-like growth factor 2 receptor	1.64	1.98
	414618	AJ204600	Hs.96978	hypothetical protein MGC10764	1.87	72.00
	414675	R79015	Hs.296281	interleukin enhancer binding factor 1	1.51	1.39
	414683	S78296	Hs.76888	hypothetical protein MGC12702	43.61	64.00
	414696	AF002020	Hs.76918	Niemann-Pick disease, type C1	28.63	71.00
55	414711	AJ310440	Hs.288735	Homo sapiens cDNA FLJ13522 fis, clone PL	14.86	42.00
	414718	H95348	Hs.107987	ESTs	1.00	5.00
	414732	AW410976	Hs.77152	minichromosome maintenance deficient (S.	1.64	1.44
	414747	U30872	Hs.77204	centromere protein F (350/400kD, mitosis	65.01	74.00
	414761	AJ077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2	130.35	121.00
60	414774	X02419	Hs.77274	plasminogen activator, urokinase	2.24	2.19
	414806	D14694	Hs.77329	phosphatidylserine synthase 1	1.63	1.53
	414809	AJ434699	Hs.77356	transferrin receptor (p90, CD71)	1.97	2.60
	414812	X72755	Hs.77367	monokine induced by gamma interferon	3.48	10.60
	414825	X06370	Hs.77432	epidermal growth factor receptor (avian	103.22	143.00
	414839	X63692	Hs.77462	DNA (cytosine-5)-methyltransferase 1	1.80	1.69
65	414883	AA926960		CDC28 protein kinase 1	14.29	10.06
	414907	X90725	Hs.77597	polo (Drosophila)-like kinase	1.95	2.20
	414914	U49844	Hs.77613	ataxia telangiectasia and Rad3 related	3.00	2.90
	414945	BE076358	Hs.77667	lymphocyte antigen 6 complex, locus E	1.02	1.21
70	414972	BE263782	Hs.77695	KIAA0008 gene product	1.00	1.00
	415014	AW954064	Hs.24951	ESTs	1.42	2.84
	415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	34.72	107.00
	415227	AW821113	Hs.72402	ESTs	1.87	49.00
	415238	R37780	Hs.21422	ESTs	1.00	1.00
75	415263	AA948033	Hs.130853	ESTs	1.00	1.00
	415295	R41450	Hs.6546	ESTs	1.00	1.00
	415339	NM_015156	Hs.78398	KIAA0071 protein	51.18	166.00
	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibito	30.84	63.00
	415674	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,	1.48	1.39
80	415709	AA649850	Hs.278558	ESTs	1.00	1.00
	415735	AA704162	Hs.120811	ESTs, Weakly similar to I38022 hypothe	1.00	72.00
	415799	AA653718	Hs.225841	DKFZP434D193 protein	6.23	31.00
	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
85	415857	AA866115	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	32.51	35.00
	415889	AJ267700		ESTs	78.89	1.00

	416018	AW138239	Hs.78977	proprotein convertase subtilisin/kexin 1	1.00	1.00
	416065	BE267931	Hs.78996	proliferating cell nuclear antigen	3.35	2.32
	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (39.03	3.00
5	416177	AA174069	Hs.187607	ESTs	1.00	9.00
	416178	AI808527	Hs.192822	serologically defined breast cancer anti	3.83	3.76
	416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
	416209	AA236776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h	9.70	1.00
	416239	AL038450	Hs.48948	ESTs	83.87	129.00
10	416250	AA581386	Hs.73452	hypothetical protein MGC10791	1.96	2.12
	416322	BE019494	Hs.79217	pyrroline-5-carboxylate reductase 1	2.08	1.73
	416423	H54375	Hs.268921	ESTs	1.00	89.00
	416448	L13210	Hs.79339	lectin, galactoside-binding, soluble, 3	1.28	1.54
	416498	U33632	Hs.79351	potassium channel, subfamily K, member 1	27.29	67.00
	416558	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
15	416661	AA634543	Hs.79440	IGF-II mRNA-binding protein 3	9.96	5.00
	416722	AA354604	Hs.122546	hypothetical protein FLJ23017	3.68	33.00
	416819	U77735	Hs.80205	pim-2 oncogene	1.59	1.84
	416836	N21352	Hs.42987	ESTs, Weakly similar to S21348 probable	1.00	1.00
20	417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
	417061	AI675944	Hs.188691	Homo sapiens cDNA FLJ12033 fis, clone HE	32.95	156.00
	417079	U65590	Hs.81134	interleukin 1 receptor antagonist	3.91	4.93
	417218	AA129547	Hs.285754	met proto-oncogene (hepatocyte growth fa	1.00	51.00
	417233	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	3.38	2.05
25	417308	H60720	Hs.81892	KIAA0101 gene product	82.94	25.35
	417315	AI080042	Hs.180450	ribosomal protein S24	106.61	121.00
	417324	AW265494		ESTs	1.20	1.28
	417366	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)	8.97	3.27
	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	2.59	1.82
30	417428	N87579	Hs.278871	gb:LL2030F Human fetal heart, Lambda ZAP	1.00	52.00
	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	304.75	173.00
	417466	AI681547	Hs.59457	hypothetical protein FLJ22127	1.24	1.34
	417512	AI979168	Hs.344096	glycoprotein (transmembrane) nmb	2.14	6.50
	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated	2.66	1.68
35	417542	J04129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
	417576	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	42.76	51.00
	417715	AW969587	Hs.86366	ESTs	6.35	2.75
	417720	AA205625	Hs.208067	ESTs	113.31	58.00
	417791	AW965339	Hs.111471	ESTs	39.98	16.00
40	417830	AW504786	Hs.122579	hypothetical protein FLJ10461	2.61	31.00
	417866	AW067903	Hs.82772	collagen, type XI, alpha 1	2.35	2.44
	417900	BE250127	Hs.82906	CDC20 (cell division cycle 20, S. cerevi	1.52	1.11
	417933	X02308	Hs.82962	thymidylate synthetase	4.74	2.55
	417944	AU077196	Hs.82985	collagen, type V, alpha 2	3.61	5.21
45	417975	AA641836	Hs.30085	hypothetical protein FLJ23186	12.49	38.00
	417991	AA731452	Hs.190008	ESTs	1.00	26.00
	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	3.02	2.12
	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	187.59	1.00
50	418054	NM_002318	Hs.83354	lysyl oxidase-like 2	2.85	2.63
	418057	NM_012151	Hs.83353	coagulation factor VIII-associated (intr	1.54	1.69
	418113	AI272141	Hs.83484	SRY (sex determining region Y)-box 4	6.82	5.22
	418140	BE613836	Hs.83551	microfibrillar-associated protein 2	1.26	1.46
	418203	X54942	Hs.83758	CDC28 protein kinase 2	134.19	144.00
	418207	C14685	Hs.34772	ESTs	1.00	1.00
55	418216	AA662240	Hs.283099	AF15q14 protein	64.66	61.00
	418236	AW994005	Hs.337534	ESTs	18.53	147.00
	418249	H89226	Hs.34892	KIAA1323 protein	30.53	106.00
	418281	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9	1.00	3.00
	418283	S79895	Hs.83942	cathepsin K (pseudosclerosis)	3.95	5.16
60	418300	AI433074	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	3.18	2.91
	418322	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	11.96	6.68
	418327	U70370	Hs.84136	paired-like homeodomain transcription fa	9.23	2.22
	418345	AJ001696	Hs.241407	serine (or cysteine) proteinase inhibito	1.00	1.00
	418379	AA218940	Hs.137516	fidgin-like 1	21.68	44.00
65	418397	NM_001269	Hs.84746	chromosome condensation 1	1.00	8.00
	418403	D86978	Hs.84790	KIAA0225 protein	16.91	18.98
	418462	BE001596	Hs.85266	Integrin, beta 4	1.58	1.16
	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	3.22	2.38
	418506	AA084248	Hs.85339	G protein-coupled receptor 39	2.66	2.22
70	418526	BE019020	Hs.85838	solute carrier family 16 (monocarboxylic	2.04	2.21
	418538	BE244323	Hs.85951	exportin, tRNA (nuclear export receptor	1.33	37.00
	418543	NM_005329	Hs.85962	hyaluronan synthase 3	1.04	1.23
	418574	N28754		M-phase phosphoprotein 9	48.60	85.00
	418592	X99226	Hs.284153	Fanconi anemia, complementation group A	18.24	26.00
75	418641	BE243136	Hs.86947	a disintegrin and metalloproteinase doma	1.19	1.41
	418661	NM_001949	Hs.1169	E2F transcription factor 3	29.05	43.00
	418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
	418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
	418686	Z36830	Hs.87268	annexin A8	1.54	1.98
80	418689	AI360883	Hs.274448	hypothetical protein FLJ11029	1.19	1.04
	418712	Z42183		gb:HSC0BF041 normalized infant brain cDN	1.00	12.00
	418727	AA227609	Hs.94834	ESTs	1.00	49.00
	418738	AW388633	Hs.5682	solute carrier family 7, (cationic amino	49.85	1.00
	418819	AA228776	Hs.191721	ESTs	1.00	140.00
85	418830	BE513731	Hs.88959	hypothetical protein MGC4816	20.97	23.00
	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	57.09	35.00

	418971	AA360392	Hs.87113	ESTs	1.00	12.00
	418973	AA233056	Hs.191518	ESTs	4.89	28.00
	419078	M93119	Hs.89584	Insulinoma-associated 1	1.00	10.00
5	419079	AW014836	Hs.18844	ESTs	1.09	1.98
	419080	AW150835	Hs.18878	hypothetical protein FLJ21620	2.06	1.68
	419088	AJ538323	Hs.52620	integrin, beta 8	15.60	51.00
	419092	J05581	Hs.89603	mucin 1, transmembrane	1.11	1.83
	419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
	419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor I	1.10	1.14
10	419183	U60669	Hs.89663	cytochrome P450, subfamily XIV (vitamin	1.00	1.00
	419216	AJ076718	Hs.164021	small inducible cytokine subfamily B (Cy	3.18	2.43
	419288	AA256106	Hs.87507	ESTs	1.00	34.00
	419335	AW960146	Hs.284137	hypothetical protein FLJ12888	1.00	8.00
15	419354	M62839	Hs.1252	apolipoprotein H (beta-2-glycoprotein I)	22.63	54.00
	419359	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)	2.50	1.98
	419423	D26488	Hs.90315	KIAA0007 protein	1.00	7.00
	419443	D62703		gb:HUM316G108 Clontech human aorta polyA	1.00	12.00
	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	1.64	1.84
20	419474	AW968619	Hs.155849	ESTs	13.63	62.00
	419485	AA489023	Hs.99807	ESTs, Weakly similar to unnamed protein	4.27	2.26
	419488	AA316241	Hs.90691	nucleophosmin/nucleoplasm 3	3.66	3.63
	419502	AJ076704		fibrinogen, A alpha polypeptide	13.05	115.00
	419539	AF070590	Hs.90869	Homo sapiens clones 24622 and 24623 mRNA	74.60	117.00
25	419556	U29615	Hs.91093	chitinase 1 (chitotriosidase)	1.47	4.98
	419569	AJ971651	Hs.91143	jagged 1 (Alagille syndrome)	1.00	4.00
	419594	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	94.30	94.00
	419703	AJ793257	Hs.128151	ESTs	15.26	50.00
	419721	NM_001650	Hs.288650	aquaporin 4	1.00	191.00
30	419729	AA586442	Hs.21411	gb:nc53a03.s1 NCL_CGAP_SS1 Homo sapiens	1.00	59.00
	419741	NM_007019	Hs.93002	ubiquitin carrier protein E2-C	2.02	1.08
	419745	AF042001	Hs.93005	slug (chicken homolog), zinc finger prot	1.00	1.00
	419752	AA249573	Hs.152618	ESTs, Moderately similar to ZN91_HUMAN Z	29.87	77.00
	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	50.99	214.00
35	419836	AJ792788		gb:at91d05.y5 NCL_CGAP_Kd5 Homo sapiens	1.00	1.00
	419937	AB040959	Hs.93836	DKFZP434N014 protein	1.64	2.47
	419983	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f	15.72	94.00
	420005	AW271108	Hs.133294	ESTs	3.15	1.43
	420047	AA78658	Hs.94631	brefeldin A-inhibited guanine nucleotide	12.45	39.00
40	420058	AK001423	Hs.94694	Homo sapiens cDNA FLJ10561 fis, clone NT	1.00	117.00
	420162	BE378432	Hs.95577	cyclin-dependent kinase 4	1.43	1.21
	420251	AW374968	Hs.348112	Human DNA sequence from clone RPS-1103G7	2.35	3.23
	420259	AF004884	Hs.96253	calcium channel, voltage-dependent, P/Q	0.77	1.15
	420281	AI623693	Hs.323494	ESTs	45.04	54.00
45	420309	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S	49.22	31.00
	420332	NM_001756	Hs.1305	serine (or cysteine) proteinase inhibitor	0.05	2.82
	420380	AA640891	Hs.102406	ESTs	0.99	2.74
	420462	AF050147	Hs.97932	chondromodulin I precursor	1.00	1.00
	420520	AK001978	Hs.99510	similar to rab11-binding protein	49.74	133.00
50	420552	AK000492	Hs.99806	hypothetical protein	94.85	88.00
	420560	AW207748	Hs.59115	ESTs	1.00	17.00
	420610	AI683183	Hs.99348	distal-less homeo box 5	1.00	13.00
	420689	H79979	Hs.88678	ESTs	50.09	95.00
	420721	AA927802	Hs.159471	ZAP3 protein	1.00	31.00
55	420759	T11832	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	1.00	48.00
	420783	AI659638	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
	420900	AL045633	Hs.44269	ESTs	2.24	7.00
	420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
	421002	AF116030	Hs.100932	transcription factor 17	1.00	27.00
60	421027	AA761198	Hs.55254	ESTs	2.87	38.00
	421037	AI684808	Hs.197653	ESTs	1.00	46.00
	421041	N36914	Hs.14691	ESTs, Moderately similar to I38022 hypot	1.00	98.00
	421073	NM_004689	Hs.101448	metastasis associated 1	1.34	1.46
	421110	AJ250717	Hs.1355	cathepsin E	119.47	427.00
65	421133	AA401369	Hs.190721	ESTs	1.10	17.00
	421150	AJ913562	Hs.189902	ESTs	1.45	1.63
	421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
	421307	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f	1.37	1.10
	421316	AA287203	Hs.324728	SMA5	1.00	21.00
70	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	1.92	3.94
	421451	AA291377	Hs.50831	ESTs	5.89	14.00
	421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.78
	421506	BE302796	Hs.105097	thymidine kinase 1, soluble	1.58	1.08
	421508	NM_004833	Hs.105115	absent in melanoma 2	5.11	5.23
75	421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase I, I	1.00	3.00
	421524	AA312082	Hs.105445	GNDF family receptor alpha 1	2.63	10.58
	421526	AL080121	Hs.105460	DKFZP564O0823 protein	1.46	1.88
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	30.21	50.32
	421574	AJ000152	Hs.105924	defensin, beta 2	1.67	1.74
80	421582	AJ910275		trifoli factor 1 (breast cancer, estroge	1.23	1.00
	421633	AF121860	Hs.106260	sorting nexin 10	1.00	116.00
	421659	NM_014459	Hs.106511	protocadherin 17	0.05	6.33
	421677	H64092	Hs.38282	ESTs	1.31	1.42
	421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	1.41	1.20
85	421773	W69233	Hs.112457	ESTs	1.12	1.14
	421777	BE562088	Hs.108196	HSPO37 protein	1.97	1.29

	421800	AA298151	Hs.222969	ESTs	1.03	1.30
	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.88	1.59
	421896	N62293	Hs.45107	ESTs	11.84	22.80
5	421928	AF013758	Hs.109643	polyadenylate binding protein-interactin	45.89	90.00
	421931	NM_000814	Hs.1440	gamma-aminobutyric acid (GABA) A recepto	1.13	1.49
	421948	L42583	Hs.334309	keratin 6A	51.83	20.25
	421975	AW961017	Hs.6459	hypothetical protein FLJ11856	1.17	1.15
	422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
10	422094	AF129535	Hs.272027	F-box only protein 5	67.61	62.00
	422095	AI668872	Hs.282604	hypothetical protein FLJ22704	4.37	2.34
	422109	S73265	Hs.1473	gastrin-releasing peptide	4.18	95.50
	422128	AW881145		gb:QV0-OT0033-010400-182-a07 OT0033 Homo	40.89	71.00
	422129	AU076635	Hs.1478	serine (or cysteine) proteinase inhibitor	1.13	1.38
15	422134	AW179019	Hs.112110	mitochondrial ribosomal protein L42	41.59	96.00
	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias	3.29	1.68
	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6	4.93	5.73
	422282	AF019225	Hs.114309	apolipoprotein L	1.49	1.71
20	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	25.99	10.91
	422310	AA316622	Hs.98370	cytochrome P450, subfamily IIS, polypept	1.54	1.41
	422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
	422330	D30783	Hs.115263	epiregulin	1.00	112.00
	422364	AF067800	Hs.115515	C-type (calcium dependent, carbohydrate-	9.39	60.00
25	422406	AF025441	Hs.116206	Opa-interacting protein 5	18.33	53.00
	422424	AJ186431	Hs.296638	prostate differentiation factor	1.71	3.21
	422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
	422487	AJ010901	Hs.198267	mucin 4, tracheobronchial	73.68	35.54
	422511	AU076442	Hs.117938	collagen, type XVII, alpha 1	173.97	26.00
30	422515	AW500470	Hs.117950	multifunctional polypeptide similar to S	4.68	2.92
	422656	AJ870435	Hs.1569	UIM homeobox protein 2	1.00	1.00
	422737	M26939	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	3.89	4.55
	422756	AA441787	Hs.119689	glycoprotein hormones, alpha polypeptide	1.05	1.46
	422765	AA409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	3.88	1.53
35	422809	AK001379	Hs.121028	hypothetical protein FLJ10549	99.56	53.00
	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
	422938	NM_001809	Hs.1594	centromere protein A (17kD)	70.46	61.00
	422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	77.74	3.00
	422960	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	5.88	8.55
40	422963	AA401369	Hs.190721	ESTs	171.41	17.00
	422976	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e	2.12	1.62
	422981	AF026445	Hs.122752	TATA box binding protein (TBP)-associate	10.49	35.00
	422986	AA319777	Hs.221974	ESTs	12.40	32.47
	423034	AL119930		gb:DKFZp761A092_r1 761 (synonym: hamy2)	16.41	60.00
45	423049	X59373	Hs.188023	ESTs, Moderately similar to HXDA_HUMAN H	1.00	1.00
	423081	AF262992	Hs.123159	sperm associated antigen 4	1.82	2.96
	423184	NM_004428	Hs.1624	ephrin-A1	1.14	1.53
	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermalys	2.14	1.69
50	423248	AA380177	Hs.125845	ribulose-5-phosphate-3-epimerase	7.18	14.00
	423309	BE006775	Hs.125782	sushi-repeat protein	21.90	64.00
	423361	AW170055	Hs.47628	ESTs	1.00	1.00
	423453	AW450737	Hs.128791	CGI-09 protein	55.52	66.00
	423511	AF036329	Hs.129715	gonadotropin-releasing hormone 2	0.88	1.17
	423516	AB007933	Hs.129729	ligand of neuronal nitric oxide synthase	1.76	5.40
55	423551	AA327598	Hs.233785	ESTs	3.54	4.33
	423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
	423575	C18863	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	38.88	70.00
	423624	AJ807408	Hs.166368	ESTs	1.00	67.00
	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
60	423642	AW452650	Hs.157148	hypothetical protein MGC13204	19.14	58.00
	423662	AA642452	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	3.61	13.57
	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	240.73	40.00
	423698	AA329796	Hs.1098	DKFZp434J1813 protein	1.00	59.00
	423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
65	423761	NM_006194	Hs.132576	paired box gene 9	1.00	1.00
	423787	AJ295745	Hs.236204	nuclear pore complex protein	7.18	6.64
	423816	AF151064		hypothetical protein	1.00	44.00
	423826	U20325	Hs.1707	cocaine- and amphetamine-regulated trans	1.00	1.00
	423849	AL157425	Hs.133315	Homo sapiens mRNA: cDNA DKFZp761J1324 (f	1.00	1.00
70	423887	AL080207	Hs.134585	DKFZP434G232 protein	1.00	1.00
	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f	31.33	31.00
	423954	AW753164	Hs.288604	KIAA1632 protein	5.81	10.87
	423961	D13685	Hs.136348	osteoblast specific factor 2 (fasciclin	3.55	3.30
	424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	233.42	68.00
75	424016	AW163729	Hs.6140	hypothetical protein MGC15730	0.93	1.01
	424028	AF055084	Hs.153692	Homo sapiens cDNA FLJ14354 fis, clone Y7	21.30	52.00
	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor	1.00	1.00
	424086	AJ351010	Hs.102267	lysyl oxidase	21.91	70.00
	424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
80	424120	T80579	Hs.290270	ESTs	1.00	1.00
	424165	AW582904	Hs.142255	Islet amyloid polypeptide	1.00	34.00
	424200	AA337221		gb:EST41944 Endometrial tumor Homo sapie	13.08	48.00
	424279	L29306	Hs.171814	tryptophan hydroxylase (tryptophan 5-mon	1.00	1.00
	424308	AW975531	Hs.154443	minichromosome maintenance deficient (S.	164.58	87.00
85	424326	NM_014479	Hs.145296	disintegrin protease	53.72	302.00
	424340	AA339036	Hs.7033	ESTs	0.83	1.15

	424351	BE622117	Hs.145567	hypothetical protein	0.93	1.03
	424364	AW383226	Hs.201189	ESTs, Weakly similar to G01763 atrophin-	7.02	3.24
	424381	AA285249	Hs.146329	protein kinase Ctk2	95.55	92.00
	424411	NM_005209	Hs.146549	crystallin, beta A2	1.63	3.25
5	424420	BE614743	Hs.146688	prostaglandin E synthase	1.63	1.33
	424441	X14850	Hs.147097	H2A histone family, member X	1.82	1.29
	424502	AF242388	Hs.149585	lansin	1.00	1.00
	424503	X06256	Hs.149609	Integrin, alpha 5 (fibronectin receptor,	1.02	2.24
	424513	BE385864	Hs.149894	mitochondrial translational initiation f	1.00	17.00
10	424539	L02911	Hs.150402	Activin A receptor, type 1 (ACVR1) (ALK	32.46	108.00
	424568	AF005418	Hs.150595	cytochrome P450, subfamily XXVIA, polype.	3.40	2.58
	424602	AK002055	Hs.151046	hypothetical protein FLJ11193	31.87	25.00
	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub	3.58	2.37
	424645	NM_014682	Hs.151449	KIAA0535 gene product	1.00	1.00
15	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	2.12	2.23
	424717	AW992292	Hs.152213	wingless-type MMTV integration site fami	1.00	1.00
	424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
	424840	D79987	Hs.153479	extra spindle poles, S. cerevisiae, homo	2.65	1.30
	424867	AJ024860	Hs.153591	Not56 (D. melanogaster)-like protein	1.23	1.05
20	424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
	424979	D87989	Hs.154073	UDP-galactose transporter related	1.35	1.35
	424999	AW953120		gb:EST365190 MAGE resequences, MAGB Homo	1.24	1.41
	425048	H05458	Hs.164502	ESTs	1.00	11.00
	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	7.46	87.00
25	425081	X74794	Hs.154443	minichromosome maintenance deficient (S.	2.52	3.82
	425118	AU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase	4.84	4.03
	425159	NM_004341	Hs.154868	carbamoyl-phosphate synthetase 2, aspart	3.62	2.73
	425202	AW962282	Hs.152049	ESTs, Weakly similar to I38022 hypotheti	1.00	53.00
	425234	AW152225	Hs.165909	ESTs, Weakly similar to I38022 hypotheti	100.77	44.00
30	425236	AW067800	Hs.155223	stanniocalcin 2	3.30	2.90
	425245	AJ751768	Hs.155314	KIAA0095 gene product	1.91	2.32
	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	1.41	1.49
	425266	J00077	Hs.155421	alpha-fetoprotein	1.00	68.00
	425274	BE281191	Hs.155462	minichromosome maintenance deficient (mi	1.97	1.63
35	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	141.49	123.00
	425349	AA425234	Hs.79886	ribose 5-phosphate isomerase A (ribose 5	1.00	84.00
	425371	D49441	Hs.155981	mesothelin	0.87	1.59
	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	14.90	5.76
	425420	BE536911	Hs.234545	hypothetical protein NUF2R	1.00	1.00
40	425424	NM_004954	Hs.157199	ELKL motif kinase	10.58	9.74
	425483	AF231022	Hs.158159	FAT tumor suppressor (Drosophila) homolo	1.74	1.40
	425566	AW162943	Hs.250618	UL16 binding protein 2	1.49	1.14
	425580	L11144	Hs.1907	galanin	53.29	233.00
	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
45	425692	D90041	Hs.155956	N-acetyltransferase 1 (arylamine N-acety	1.00	55.00
	425695	NM_005401	Hs.159238	protein tyrosine phosphatase, non-recept	1.00	10.00
	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.00	41.00
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	1.00	48.00
	425810	AJ923627	Hs.31903	ESTs	27.39	98.00
50	425811	AL039104	Hs.159557	karyopherin alpha 2 (RAG cohort 1, impor	1.99	1.58
	425849	AJ077288	Hs.296323	serum/glucocorticoid regulated kinase	71.16	3.42
	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	1.35	1.34
	426067	AA401369	Hs.190721	ESTs	1.01	17.00
	426088	AF038007	Hs.166195	ATPase, Class I, type 8B, member 1	26.26	47.00
55	426215	AW067800	Hs.155223	stanniocalcin 2	1.91	2.90
	426227	U67058	Hs.154299	Human proteinase activated receptor-2 mR	22.40	25.00
	426269	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f	1.00	1.00
	426283	NM_003937	Hs.169139	kynureninase (L-kynurenine hydrolase)	91.39	229.00
	426329	AL389951	Hs.271623	nucleoporin 50kD	4.34	4.08
60	426427	M86699	Hs.169840	TTK protein kinase	7.02	1.00
	426432	AF001601	Hs.169857	paraoxonase 2	1.16	1.68
	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu	2.59	1.71
	426459	AF151812	Hs.169992	hypothetical 43.2 Kd protein	1.56	1.66
	426471	M22440	Hs.170009	transforming growth factor, alpha	20.60	26.00
65	426496	D31765	Hs.170114	KIAA0061 protein	9.81	22.00
	426501	AA401369	Hs.190721	ESTs	19.23	17.00
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	103.74	41.00
	426536	AJ949749	Hs.44441	ESTs	4.65	23.00
	426572	AB037783	Hs.170623	hypothetical protein FLJ11183	1.00	43.00
70	426682	AV660038	Hs.2056	UDP glycosyltransferase 1 family, polype	160.06	8.00
	426691	NM_006201	Hs.171834	PCTAIRE protein kinase 1	1.51	1.35
	426746	J03626	Hs.2057	uridine monophosphate synthetase (orotid	2.13	1.68
	426752	X69490	Hs.172004	tin	0.02	5.14
	426784	U03749	Hs.172216	chromogranin A (parathyroid secretory pr	1.72	1.71
75	426807	AA385315	Hs.156682	ESTs	1.30	1.64
	426812	AF105365	Hs.172613	solute carrier family 12 (potassium/chlo	1.47	1.53
	426814	AF036943	Hs.172619	myelin transcription factor 1-like	1.00	1.00
	426831	BE296216	Hs.172673	S-adenosylhomocysteine hydrolase	1.51	1.25
	426897	AA401369	Hs.190721	ESTs	141.56	17.00
80	426925	NM_001196	Hs.315589	Homo sapiens cDNA: FLJ22373 fis, clone H	32.61	38.00
	426935	NM_000088	Hs.172928	collagen, type I, alpha 1	2.65	3.16
	426964	AA393739	Hs.287416	Homo sapiens cDNA FLJ11439 fis, clone HE	1.97	3.49
	426966	AJ493134		sclerostin	1.00	1.00
	426991	AK001536		Homo sapiens cDNA FLJ10574 fis, clone NT	3.39	2.28
85	427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00

	427239	BE270447	Hs.174070	ubiquitin carrier protein	1.58	1.05
	427260	AA663848		gb:ae70b06.s1 Stratagene schizo brain S1	1.34	1.60
	427281	AA906147	Hs.102869	ESTs	1.00	66.00
5	427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
	427354	T57895	Hs.191095	ESTs	1.17	1.95
	427356	AW023482	Hs.97849	ESTs	7.31	41.00
	427376	AA401533	Hs.19440	ESTs	1.00	57.00
	427383	NM_005411	Hs.177582	surfactant, pulmonary-associated protein	0.42	1.32
10	427427	AF077345	Hs.177936	lectin, superfamily member 1 (cartilage-	1.00	20.00
	427441	AA412605	Hs.343879	SPANX family, member C	1.00	1.00
	427445	X80818	Hs.178078	glutamate receptor, metabotropic 4	0.97	1.03
	427505	AA361562	Hs.178761	26S proteasome-associated pad1 homolog	4.60	4.04
	427510	Z47542	Hs.179312	small nuclear RNA activating complex, po	22.00	45.00
15	427528	AU077143	Hs.179565	minichromosome maintenance deficient (S.	97.45	92.00
	427546	AA188763	Hs.36793	hypothetical protein FLJ23188	1.50	3.24
	427562	R56424	Hs.26534	ESTs	6.81	40.00
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	69.91	62.00
	427660	AJ741320	Hs.114121	Homo sapiens cDNA: FLJ23228 fis, clone C	2.70	49.00
20	427666	AJ791495	Hs.180142	calmodulin-like skin protein	1.37	1.88
	427668	AA298760	Hs.180191	hypothetical protein FLJ14904	29.55	67.00
	427677	NM_007045	Hs.180256	FGFR1 oncogene partner	3.52	2.63
	427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
	427711	M31659	Hs.180408	solute carrier family 25 (mitochondrial	15.84	70.00
	427719	AJ393122	Hs.134726	ESTs	7.03	4.52
25	427722	AK000123	Hs.180479	hypothetical protein FLJ20116	2.92	1.74
	427747	AW411425	Hs.180655	serine/threonine kinase 12	1.76	1.28
	427912	AL022310	Hs.181097	tumor necrosis factor (ligand) superfam	9.63	59.00
	427961	AW293165	Hs.143134	ESTs	41.97	118.00
30	428004	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	23.82	1.00
	428023	AL038843		Homo sapiens cDNA: FLJ23602 fis, clone L	1.40	1.33
	428046	AW812795	Hs.337534	ESTs, Moderately similar to I38022 hypot	96.28	167.00
	428093	AW594506	Hs.104830	ESTs	1.25	1.29
	428098	AJ077258	Hs.182429	protein disulfide isomerase-related prot	1.86	1.60
35	428129	AJ244311	Hs.26912	ESTs	1.00	42.00
	428169	AJ289884	Hs.182793	golgi phosphoprotein 2	2.76	2.11
	428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	85.59	181.00
	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	8.57	21.64
40	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	7.77	15.90
	428434	AJ909935	Hs.65551	Homo sapiens, Similar to DNA segment, Ch	0.58	1.43
	428450	NM_014791	Hs.184339	KIAA0175 gene product	237.53	204.00
	428471	X57348	Hs.184510	stratifin	6.00	4.60
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	56.54	16.00
45	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	3.53	2.15
	428506	AL035461	Hs.2281	chromogranin B (secretogranin 1)	1.00	1.00
	428532	AF157326	Hs.184786	TBP-interacting protein	1.00	58.00
	428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	1.00	1.00
50	428698	AA852773	Hs.334838	KIAA1866 protein	187.37	255.00
	428728	NM_016625	Hs.191381	hypothetical protein	47.24	80.00
	428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
	428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
	428771	AB028992	Hs.193143	KIAA1069 protein	1.98	92.00
55	428801	AW277121	Hs.254881	ESTs	1.67	6.15
	428810	AF068236	Hs.193788	nitric oxide synthase 2A (inducible, hep	1.03	1.27
	428839	AJ767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	124.17	43.00
	428845	AL157579	Hs.153610	KIAA0751 gene product	1.00	1.00
	428959	AF100778	Hs.194680	WNT1 inducible signaling pathway protein	15.16	27.00
60	428969	AF120274	Hs.194699	artemin	1.36	1.24
	429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like	0.97	3.31
	429065	AJ753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	6.82	16.47
	429164	AJ688663	Hs.116586	ESTs	19.08	67.00
	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
65	429183	AB014604	Hs.197955	KIAA0704 protein	79.72	104.00
	429201	X03178	Hs.198246	group-specific component (vitamin D bind	1.00	1.00
	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	1.33	1.09
	429220	AW207206		ESTs	1.00	7.00
	429228	AJ553633	Hs.326447	ESTs	39.47	29.25
70	429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
	429276	AF056085	Hs.198612	G protein-coupled receptor 51	3.70	142.00
	429359	W00482	Hs.2399	matrix metalloproteinase 14 (membrane-in	1.30	1.94
	429412	NM_006235	Hs.2407	POU domain, class 2, associating factor	94.09	86.00
75	429413	NM_014058	Hs.201877	DESC1 protein	41.91	10.00
	429486	AF155827	Hs.203953	hypothetical protein FLJ10339	12.19	1.00
	429504	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	1.61	1.08
	429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
	429547	AA401369	Hs.190721	ESTs	1.06	17.00
80	429551	AW450624	Hs.220931	ESTs	2.89	65.00
	429563	BE619413	Hs.2437	eukaryotic translation initiation factor	1.49	1.37
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	61.86	100.00
	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	1.59	1.69
	429612	AF062649	Hs.252587	plutary tumor-transforming 1	2.78	1.74
85	429616	AJ982722	Hs.120845	ESTs	1.00	1.00
	429656	X05608	Hs.211584	neurofilament, light polypeptide (68kD)	1.00	4.00

	429663	M58874	Hs.211587	phospholipase A2, group IVA (cytosolic,	69.95	104.00
	429736	AF125304	Hs.212680	tumor necrosis factor receptor superfam	1.25	1.21
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	1.00	7.00
	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	11.80	1.00
5	429918	AW873986	Hs.119383	ESTs	1.00	78.00
	429978	AA249027		ribosomal protein S6	1.98	3.09
	429986	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	1.00	48.00
	430044	AA464510	Hs.152812	ESTs	69.27	59.00
10	430114	AA847744	Hs.99640	ESTs	1.00	1.00
	430134	BE380149	Hs.105223	ESTs, Weakly similar to T33188 hypotheti	1.00	51.00
	430147	R60704	Hs.234434	hairy/enhancer-of-split related with YRP	1.10	2.22
	430287	AW182459	Hs.125759	ESTs, Weakly similar to LEU5_HUMAN LEUKE	1.00	127.00
	430294	AI538226	Hs.32976	guanine nucleotide binding protein 4	3.80	1.47
15	430300	U60805	Hs.238648	oncostatin M receptor	1.00	35.00
	430315	NM_004293	Hs.239147	guanine deaminase	92.31	28.00
	430337	M36707	Hs.239600	catmodulin-like 3	1.18	1.08
	430378	Z29572	Hs.2556	tumor necrosis factor receptor superfam	5.28	66.00
	430388	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	16.76	38.00
20	430393	BE185030	Hs.241305	estrogen-responsive B box protein	1.63	1.50
	430439	AL133561		DKFZP434B061 protein	1.00	1.00
	430451	AA836472	Hs.297939	cathepsin B	1.64	2.12
	430454	AW469011	Hs.105635	ESTs	63.35	44.00
	430466	AF052573	Hs.241517	polymerase (DNA directed), theta	2.47	1.91
25	430481	AA479578	Hs.203269	ESTs, Moderately similar to ALU8_HUMAN A	1.00	31.00
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
	430508	AI015435	Hs.104637	ESTs	4.75	7.27
	430533	AA480895	Hs.57749	ESTs, Weakly similar to T17288 hypotheti	1.00	1.00
	430563	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.00	1.59
30	430677	Z26317	Hs.94560	desmoglein 2	1.72	1.30
	430678	AA401369	Hs.190721	ESTs	0.90	17.00
	430686	NM_001942	Hs.2633	desmoglein 1	1.00	1.00
	430788	AI742925	Hs.7179	ESTs, Weakly similar to 2004399A chromos	1.62	1.84
	430890	X54232	Hs.2699	glypican 1	1.58	1.40
35	430935	AW072916		zinc finger protein 131 (clone pHZ-10)	90.28	132.00
	430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
	431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
	431092	AI332764	Hs.125757	ESTs	13.46	63.00
40	431124	AF284221	Hs.69506	doublesex and mab-3 related transcriptio	49.43	62.00
	431164	AA493650	Hs.94367	Homo sapiens cDNA: FLJ23494 fis, clone L	0.44	2.20
	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	182.26	101.00
	431221	AW207837	Hs.285145	SRB7 (suppressor of RNA polymerase B, ye	4.15	13.97
	431277	AA501806	Hs.345824	ESTs	1.00	86.00
45	431322	AW970622		gb:EST382704 MAGE resequences, MAGK Homo	40.55	200.00
	431342	AW971018	Hs.21659	ESTs	1.00	53.00
	431384	BE158000	Hs.285026	gb:MR2-HT0377-150200-202-e03 HT0377 Homo	0.94	1.14
	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	1.30	1.25
	431494	AA991355	Hs.298312	hypothetical protein DKFZp434A1315	3.90	26.00
50	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	1.41	1.87
	431548	AI834273	Hs.9711	novel protein	5.66	15.00
	431630	NM_002204	Hs.265829	integrin, alpha 3 (antigen CD49C, alpha	0.99	1.44
	431745	AW972448	Hs.163425	ESTs	0.99	3.51
55	431770	BE221880	Hs.268555	5'-3' exonuclease 2	67.12	91.00
	431830	Y16645	Hs.271387	small inducible cytokine subfamily A (Cy	3.36	4.71
	431846	BE019924	Hs.271580	uropodkin 1B	4.49	2.51
	431890	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	2.20	3.32
	431934	AB031481	Hs.272214	STG protein	1.01	1.04
60	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	51.17	46.35
	432006	AL137382	Hs.272320	Homo sapiens mRNA; cDNA DKFZp434L1226 (f	0.94	1.65
	432023	R43020	Hs.236223	EST	0.94	47.00
	432201	AI538613	Hs.298241	Transmembrane protease, serine 3	1.10	2.24
	432210	AI567421	Hs.273330	Homo sapiens, clone IMAGE:3544662, mRNA,	1.42	1.45
65	432226	AW182766	Hs.273559	phosphate cytidylyltransferase 1, cholin	1.00	1.00
	432239	X81334	Hs.2936	matrix metalloproteinase 13 (collagenase	18.67	1.00
	432265	BE382679	Hs.285753	SCG10-like-protein	1.09	1.21
	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	40.98	58.00
	432365	AK001106	Hs.274419	hypothetical protein FLJ10244	1.00	214.00
70	432374	W68815	Hs.301885	Homo sapiens cDNA FLJ11346 fis, clone PL	157.34	37.00
	432375	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
	432407	AA221036		gb:zr03f12r1 Stratagene NT2 neuronal pr	73.71	75.00
	432441	AW292425	Hs.163484	ESTs	56.35	72.00
	432469	AI804855	Hs.207530	ESTs	1.00	24.00
75	432543	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 fis, clone C	137.72	98.00
	432552	AI537170	Hs.173725	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.00	31.00
	432583	AW023624	Hs.162282	potassium channel TASK-4; potassium chan	0.27	35.18
	432606	NM_002104	Hs.3066	granzyme K (serine protease, granzyme 3;	2.87	6.22
	432625	AI243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin	26.63	56.00
	432653	N62096	Hs.293185	ESTs, Weakly similar to JCT328 amino aci	1.92	5.29
80	432677	NM_004482	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	45.13	31.00
	432753	NM_014075	Hs.336938	Homo sapiens PRO0593 mRNA, complete cds	1.00	68.00
	432788	AA521091	Hs.178499	Homo sapiens cDNA: FLJ23117 fis, clone L	2.69	3.67
	432842	AW674093	Hs.334822	hypothetical protein MGC4485	1.22	1.34
85	432867	AW016936	Hs.233364	ESTs	1.00	1.00
	432917	NM_014125	Hs.241517	PRO0327 protein	10.25	6.62

	432920	U37689	Hs.3128	polymerase (RNA) II (DNA directed) polyp	1.44	1.30
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	154.79	85.64
	433023	AW864793	Hs.87409	thrombospondin 1	20.96	100.00
5	433042	AW193534	Hs.281895	Homo sapiens cDNA FLJ11660 fis, clone HE	1.00	10.00
	433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
	433159	AB035898	Hs.150587	kinesin-like protein 2	13.82	39.00
	433183	AF231338	Hs.222024	transcription factor BMAL2	1.00	69.00
	433258	AA622788	Hs.203613	ESTs, Weakly similar to ALUB_HUMAN IIII	1.00	1.25
	433409	AJ278802	Hs.25661	ESTs	44.81	117.00
10	433437	U20536	Hs.3280	caspase 6, apoptosis-related cysteine pr	70.39	105.00
	433485	AJ493076	Hs.201967	aldo-keto reductase family 1, member C2	11.55	2.00
	433537	AJ733692	Hs.112488	ESTs	8.66	55.00
	433547	W04978	Hs.303023	beta tubulin 1, class VI	25.16	83.00
	433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
15	433647	AA603367	Hs.222294	ESTs	20.30	49.00
	433658	L03678	Hs.156110	immunoglobulin kappa constant	5.92	10.03
	433800	AJ094221	Hs.135150	lung type-I cell membrane-associated gly	2.29	2.22
	433819	AW511097	Hs.112765	ESTs	3.71	8.00
	433862	D86960	Hs.3610	KIAA0205 gene product	62.08	104.00
20	433980	AA137152	Hs.286049	phosphoserine aminotransferase	108.91	47.00
	434088	AF116677	Hs.249270	hypothetical protein PRO1956	1.00	1.00
	434094	AA305599	Hs.238205	hypothetical protein PRO2013	121.27	87.00
	434105	AW952124	Hs.13094	presenilins associated rhomboid-like pro	1.22	1.23
	434217	AW014795	Hs.23349	ESTs	14.11	57.00
25	434340	AJ193043	Hs.128685	ESTs, Weakly similar to T17226 hypotheti	2.10	2.56
	434360	AA401369	Hs.190721	ESTs	40.98	17.00
	434414	AJ798376	Hs.325335	gb:tr34b07.x1 NCL CGAP_Ov23 Homo sapiens	1.48	1.56
	434424	AJ811202	Hs.231853	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
30	434467	BE562368	Hs.231853	Homo sapiens cDNA FLJ13445 fis, clone PL	54.91	85.00
	434551	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	2.46	2.00
	434627	AJ221894	Hs.39311	ESTs	1.00	1.00
	434699	AA643697	Hs.149425	Homo sapiens cDNA FLJ11980 fis, clone HE	1.00	23.00
	434769	AA648884	Hs.134278	Homo sapiens cDNA FLJ12676 fis, clone NT	7.08	56.00
	434792	AA649253	Hs.132458	ESTs	8.52	44.00
35	434808	AF155108	Hs.256150	Homo sapiens, Similar to RIKEN cDNA 2810	11.33	1.00
	434828	D90070	Hs.96	phorbol-12-myristate-13-acetate-induced	1.00	1.00
	434876	AF160477	Hs.61460	Ig superfamily receptor LNIR	1.25	1.29
	434891	AA814309	Hs.123583	ESTs	1.00	6.00
	434928	AW015595	Hs.4267	Homo sapiens clones 24714 and 24715 mRNA	1.00	1.00
40	435013	H91923	Hs.110024	Target CAT	1.26	1.10
	435066	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin	1.69	1.37
	435087	AW975241	Hs.23567	ESTs	1.00	1.00
	435099	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.90	1.93
	435159	AA668879	Hs.116649	ESTs	1.00	1.00
45	435205	X54136	Hs.181125	immunoglobulin lambda locus	1.02	1.46
	435232	NM_001262	Hs.4854	cyclin-dependent kinase inhibitor 2C (p1	2.04	2.70
	435304	H10709	Hs.269524	ESTs	27.58	139.00
	435313	AJ769400	Hs.189729	ESTs	1.00	14.00
	435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
50	435509	AJ458679	Hs.181915	ESTs	1.00	1.00
	435525	AJ831297	Hs.123310	ESTs	1.00	56.00
	435532	AW291488	Hs.117305	Homo sapiens, clone IMAGE:3682908, mRNA	1.00	2.00
	435550	AJ224456	Hs.324507	H.sapiens polyA site DNA	3.42	3.92
	435602	AF217515	Hs.283532	uncharacterized bone marrow protein BM03	3.95	1.80
55	435766	R11673	Hs.186498	ESTs	1.00	28.00
	435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
	436069	AJ056879	Hs.263209	ESTs	1.00	58.00
	436170	AW450381	Hs.14529	ESTs	1.00	18.00
60	436211	AK001581	Hs.334828	hypothetical protein FLJ10719; KIAA1794	5.84	22.00
	436213	AA325512	Hs.71472	hypothetical protein FLJ10774; KIAA1709	1.42	1.27
	436217	T53925	Hs.107	fibrinogen-like 1	57.97	31.00
	436238	AK002163	Hs.301724	hypothetical protein FLJ11301	2.51	1.71
	436251	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)	2.33	1.64
	436291	BE568452	Hs.344037	protein regulator of cytokinesis 1	108.99	52.00
65	436302	AL355841	Hs.99330	hypothetical protein FLJ23588	0.75	2.81
	436396	AW992292	Hs.152213	wingless-type MMTV integration site faml	60.01	1.00
	436414	BE264633	Hs.143638	WD repeat domain 4	2.50	2.19
	436419	AJ948626	Hs.171356	ESTs	0.95	1.33
	436443	AW138211	Hs.128746	ESTs	1.12	9.26
70	436474	AJ270693	Hs.199887	ESTs	1.00	1.00
	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.28	1.56
	436486	AA742221	Hs.120633	ESTs	1.00	19.00
	436511	AA721252	Hs.291502	ESTs	16.76	14.00
	436553	X57809	Hs.181125	immunoglobulin lambda locus	1.08	1.74
75	436557	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	19.20	9.75
	436608	AA628980	Hs.127680	down syndrome critical region protein DS	33.92	25.00
	436667	AW025183	Hs.292979	ESTs	0.89	1.19
	436771	AW975687	Hs.190721	ESTs	1.00	10.00
	436839	AA401369	Hs.190721	ESTs	1.00	17.00
80	436887	AW953157	Hs.193235	hypothetical protein DKFZp547D155	1.06	1.15
	436944	AW268814	Hs.5840	ESTs	1.00	1.00
	436961	AW375974	Hs.156704	ESTs	25.13	25.00
	436972	AA284679	Hs.25640	claudin 3	1.59	1.46
	437016	AJ076916	Hs.5398	guanine monophosphate synthetase	2.35	1.78
85	437044	AL035864	Hs.69517	cDNA for differentially expressed CO16 g	1.34	1.13

	437181	AI306615	Hs.125343	ESTs, Weakly similar to KIAA0758 protein	1.00	17.00
	437204	AL110216	Hs.22826	ESTs, Weakly similar to I55214 salivary	40.55	82.00
	437205	AL110232	Hs.27243	Homo sapiens mRNA; cDNA DKFZp554D2071 (f	1.00	112.00
5	437259	AI377755	Hs.120685	ESTs	1.00	205.00
	437270	R18087	Hs.323769	cisplatin resistance related protein CRR	1.55	1.54
	437271	AL137445	Hs.28846	Homo sapiens mRNA; cDNA DKFZp556O134 (fr	113.25	125.00
	437370	AL359567	Hs.161962	Homo sapiens mRNA; cDNA DKFZp547D023 (fr	1.82	4.57
	437390	AL125859	Hs.112607	ESTs	1.35	1.75
10	437412	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	3.58	3.20
	437435	AI306152	Hs.27027	hypothetical protein DKFZp762H1311	3.03	1.08
	437444	H46008	Hs.31518	ESTs	1.00	39.00
	437558	AI954795	Hs.156135	ESTs	1.00	19.00
	437623	D63880	Hs.5719	chromosome condensation-related SMC-asso	1.95	1.57
	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	1.00	3.00
15	437814	AI088192	Hs.135474	ESTs, Weakly similar to DDX9_HUMAN ATP-D	1.00	45.00
	437840	AA884835	Hs.292014	ESTs	1.07	1.78
	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 (Hsa	1.68	3.26
	437879	BE262082	Hs.5894	hypothetical protein FLJ10305	1.87	2.52
20	437915	AI637993	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca	74.05	35.00
	437916	BE566249	Hs.20999	hypothetical protein FLJ23142	23.15	89.00
	437937	AI917222	Hs.121655	ESTs	1.00	1.00
	437942	AI888256	Hs.307525	ESTs	12.28	31.00
	438091	AW373062		nuclear receptor subfamily 1, group I, m	1.53	10.85
25	438113	AI467908	Hs.8882	ESTs	1.80	2.39
	438119	AW963217	Hs.203961	ESTs, Moderately similar to AF116721 89	22.67	36.90
	438274	AI918906	Hs.55080	ESTs	1.00	1.00
	438378	AW970529	Hs.86434	hypothetical protein FLJ21816	38.92	38.00
	438403	AA806607	Hs.292206	ESTs	1.00	1.00
30	438494	AA908678	Hs.130183	ESTs	2.05	80.00
	438546	AW297204	Hs.125811	ESTs	1.00	131.00
	438552	AJ245820	Hs.6314	type I transmembrane receptor (seizure-r	1.43	1.45
	438702	AI879064	Hs.54618	ESTs	1.00	34.00
	438724	AW612553	Hs.114670	Human DNA sequence from clone RP11-16L21	1.33	1.10
35	438746	AI885815	Hs.184727	Human melanoma-associated antigen p97 (m	2.42	1.59
	438779	NM_003787	Hs.6414	nucleolar protein 4	1.00	18.00
	438821	AA826425	Hs.192375	ESTs	2.03	2.57
	438885	AI866558	Hs.184987	ESTs	6.42	88.00
	438898	AA401369	Hs.190721	ESTs	22.41	17.00
40	438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	1.00	1.00
	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-85O9	2.20	1.88
	439000	AW979121		gb:EST391231 MAGE resequences, MAGP Homo	2.78	4.81
	439023	AA745978	Hs.28273	ESTs	1.17	1.31
	439024	R96696	Hs.35598	ESTs	1.00	28.00
45	439128	AI949371	Hs.153089	ESTs	1.00	67.00
	439146	AW138909	Hs.156110	immunoglobulin kappa constant	1.38	1.41
	439223	AW238299	Hs.250618	UL16 binding protein 2	1.93	1.64
	439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
	439318	AW837046	Hs.6527	G protein-coupled receptor 56	2.00	2.20
50	439343	AF086161	Hs.114611	hypothetical protein FLJ11808	6.10	7.37
	439394	AA401369	Hs.190721	ESTs	3.39	17.00
	439410	AA632012	Hs.188745	ESTs	1.83	3.07
	439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00
	439452	AA918317	Hs.67987	B-cell CLL/lymphoma 11B (zinc finger pro	18.76	122.00
55	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13	2.78	1.58
	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	1.22	1.44
	439492	AF086310	Hs.103159	ESTs	7.46	39.00
	439523	W72348	Hs.185029	ESTs	1.00	1.19
	439592	AF086413	Hs.58399	ESTs	1.00	1.00
60	439606	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
	439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
	439702	AW085525	Hs.134182	ESTs	4.30	10.00
	439706	AW872527	Hs.59761	ESTs, Weakly similar to DAPI1_HUMAN DEATH	86.55	11.00
	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	2.35	1.88
65	439750	AL359053	Hs.57664	Homo sapiens mRNA full length insert cDN	2.02	6.08
	439759	AL359055	Hs.67709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	439780	AL109688		gb:Homo sapiens mRNA full length insert	7.27	25.00
	439840	AW449211	Hs.105445	GDNF family receptor alpha 1	1.00	1.00
	439926	AW014875	Hs.137007	ESTs	32.58	71.00
70	439963	AW247529	Hs.6793	platelet-activating factor acetylhydrola	21.28	9.55
	439979	AW600291	Hs.6823	hypothetical protein FLJ10430	68.83	61.00
	440006	AK000517	Hs.6844	hypothetical protein FLJ20510	1.83	4.02
	440028	AW473675	Hs.125843	ESTs, Weakly similar to T17227 hypotheti	1.42	2.54
	440106	AA854958	Hs.127699	KIAA1603 protein	1.00	54.00
75	440138	AB033023	Hs.318127	hypothetical protein FLJ10201	24.18	52.00
	440273	AI805392	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	3.21	4.72
	440289	AW450991	Hs.192071	ESTs	38.63	113.00
	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
	440492	R39127	Hs.21433	hypothetical protein DKFZp547J036	2.35	3.62
80	440527	AV657117	Hs.184164	ESTs, Moderately similar to S65657 alpha	10.84	57.00
	440659	AF134160	Hs.7327	claudin 1	3.18	2.37
	440704	M69241	Hs.162	insulin-like growth factor binding prote	2.89	2.09
	440943	AW082298	Hs.146161	hypothetical protein MGC2408	2.02	1.41
	440994	AI160011	Hs.272068	ESTs	1.29	1.14
85	441020	AA401369	Hs.190721	ESTs	142.99	17.00
	441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00

	441128	AA570256	ESTs, Weakly similar to T23273 hypotheti	4.13	3.50
	441290	W27501	Hs.89605 cholinergic receptor, nicotinic, alpha p	1.00	1.00
	441352	BE614410	Hs.23044 RAD51 (S. cerevisiae) homolog (E coli Ra	130.23	43.00
	441377	BE218239	Hs.202656 ESTs	22.03	1.00
5	441390	AI692560	Hs.131175 ESTs	3.65	7.70
	441497	R51064	Hs.23172 ESTs	1.00	1.00
	441525	AW241867	Hs.127728 ESTs	1.53	1.42
	441553	AA281219	Hs.121296 ESTs	1.89	1.57
	441607	NM_005010	Hs.7912 neuronal cell adhesion molecule	1.47	2.11
10	441633	AW958544	Hs.112242 normal mucosa of esophagus specific 1	216.22	363.00
	441636	AA081846	Hs.7921 Homo sapiens mRNA; cDNA DKFZp566E183 (fr	2.31	2.05
	441737	X79449	Hs.7957 adenosine deaminase, RNA-specific	1.30	1.49
	441790	AA401369	Hs.190721 ESTs	44.15	17.00
	441801	AW242799	Hs.86366 ESTs	1.00	1.00
15	441919	AI553802	Hs.128121 ESTs	1.00	122.00
	441937	R41782	Hs.22279 ESTs	0.86	1.37
	441954	AI744935	Hs.8047 Fanconi anemia, complementation group G	1.48	1.39
	442025	AW887434	Hs.11810 CDA11 protein	1.00	46.00
	442029	AW956698	Hs.14456 neural precursor cell expressed, develop	9.92	45.00
20	442072	AI740832	Hs.12311 Homo sapiens clone 23570 mRNA sequence	25.05	77.00
	442108	AW452649	Hs.166314 ESTs	3.61	3.14
	442117	AW664964	Hs.128899 ESTs	3.00	5.49
	442137	AA977235	Hs.128830 ESTs, Weakly similar to Z192_HUMAN ZINC	1.00	1.00
	442159	AW163390	Hs.278554 heterochromatin-like protein 1	1.92	1.66
25	442179	AA983842	Hs.333555 chromosome 2 open reading frame 2	27.22	50.00
	442328	AI952430	Hs.150614 ESTs, Weakly similar to ALU4_HUMAN ALU S	5.00	3.42
	442432	BE093589	Hs.38178 hypothetical protein FLJ123468	181.59	76.00
	442530	AI580830	Hs.176508 Homo sapiens cDNA FLJ14712 fis, clone NT	10.59	144.00
	442547	AA306997	Hs.217484 ESTs, Weakly similar to ALU1_HUMAN ALU S	109.23	98.00
30	442556	AI137761	Hs.8379 Homo sapiens mRNA; cDNA DKFZp586L2424 (f	1.00	53.00
	442619	AA447492	Hs.20183 ESTs, Weakly similar to AF164793 1 prote	29.02	50.00
	442710	AI015631	Hs.23210 ESTs	1.00	19.00
	442717	R88362	Hs.180591 ESTs, Weakly similar to T23976 hypotheti	1.00	5.00
35	442875	BE623003	Hs.23625 Homo sapiens clone TCCCTA00142 mRNA sequ	22.85	50.00
	442914	AW188551	Hs.99519 hypothetical protein FLJ14007	25.33	82.00
	442932	AA457211	Hs.8858 bromodomain adjacent to zinc finger doma	3.18	4.41
	442942	AW167087	Hs.131562 ESTs	8.45	64.00
	443068	AI188710	ESTs	1.00	27.00
40	443204	AW205876	Hs.29643 Homo sapiens cDNA FLJ13103 fis, clone NT	1.00	24.00
	443211	AI128388	Hs.143655 ESTs	12.42	2.00
	443247	BE614387	Hs.333893 c-Myc target JPO1	128.84	96.00
	443324	R44013	Hs.164225 ESTs	0.02	4.59
	443383	AI792453	Hs.166507 ESTs	1.00	47.00
	443400	R28424	Hs.250648 ESTs	18.52	61.00
45	443426	AF098158	Hs.9329 chromosome 20 open reading frame 1	4.02	1.75
	443572	AA025610	Hs.9605 cleavage and polyadenylation specific fa	2.98	2.57
	443575	AI078022	Hs.269636 ESTs, Weakly similar to ALU1_HUMAN ALU S	1.00	29.00
	443614	AV655386	Hs.7645 fibrinogen, B beta polypeptide	1.00	16.00
	443633	AL031290	Hs.9654 similar to pregnancy-associated plasma p	1.00	39.00
50	443648	AI085377	Hs.143610 ESTs	39.81	70.00
	443715	AI583187	Hs.9700 cyclin E1	48.74	7.00
	443723	AI144442	Hs.157144 syntaxin 6	1.29	1.30
	443802	AW504924	Hs.9805 KIAA1291 protein	1.75	1.61
55	443859	NM_013409	Hs.9914 follistatin	1.35	1.13
	443892	AA401369	Hs.190721 ESTs	1.00	17.00
	443947	W24187	gbzb47f09.r1 Soares_fetal_lung_NbHL19W	1.33	1.64
	443991	NM_002250	Hs.10082 potassium intermediate/small conductance	5.71	6.87
	444006	BE395085	Hs.10086 type I transmembrane protein Fn14	1.47	1.92
60	444009	AI380792	Hs.135104 ESTs	1.00	77.00
	444017	U04840	Hs.214 neuro-oncological ventral antigen 1	1.00	1.00
	444127	N63620	Hs.13281 ESTs	1.00	29.00
	444129	AW294292	Hs.256212 ESTs	1.00	1.00
	444279	U62432	Hs.89605 cholinergic receptor, nicotinic, alpha p	0.60	7.80
65	444371	BE540274	Hs.239 forkhead box M1	2.91	1.14
	444378	R41339	Hs.12569 ESTs	1.00	1.00
	444381	BE387335	Hs.283713 ESTs, Weakly similar to S64054 hypotheti	469.00	556.00
	444461	R53734	Hs.25978 ESTs, Weakly similar to 2109260A B cell	12.88	105.00
	444471	AB020684	Hs.11217 KIAA0877 protein	24.91	90.00
70	444489	AI151010	Hs.157774 ESTs	1.00	111.00
	444619	BE538082	Hs.8172 ESTs, Moderately similar to A46010 X-in	1.00	70.00
	444665	BE613126	Hs.47783 B aggressive lymphoma gene	30.56	139.00
	444707	AI188613	Hs.41690 desmocollin 3	1.00	1.00
	444735	BE019923	Hs.243122 hypothetical protein FLJ13057 similar to	77.02	90.00
75	444781	NM_014400	Hs.11950 GPI-anchored metastasis-associated prote	1.57	1.31
	444783	AK001468	Hs.62180 anillin (Drosophila Scraps homolog), act	77.55	2.00
	445236	AK001676	Hs.12457 hypothetical protein FLJ10814	1.00	27.00
	445258	AI635931	Hs.147613 ESTs	1.00	73.00
	445413	AA151342	Hs.12577 CG-147 protein	28.14	50.00
80	445417	AK001058	Hs.12680 Homo sapiens cDNA FLJ10196 fis, clone HE	1.81	2.62
	445443	AV653838	Hs.322971 ESTs	1.00	1.00
	445462	AA378776	Hs.288649 hypothetical protein MGC3077	2.09	1.70
	445517	AF208855	Hs.12830 hypothetical protein	1.67	70.00
	445537	AI245671	Hs.12844 EGF-like-domain, multiple 6	1.71	2.72
85	445580	AF167572	Hs.12912 skb1 (S. pombe) homolog	1.52	1.34
	445654	X91247	Hs.13046 thioredoxin reductase 1	1.51	1.52

	445669	AI570830	Hs.174870	ESTs	10.95	11.45
	445818	BE045321	Hs.136017	ESTs	1.00	1.00
	445873	AA250970	Hs.251946	poly(A)-binding protein, cytoplasmic 1-I	49.42	54.00
5	445885	AI734009	Hs.127699	KIAA1603 protein	1.00	132.00
	445898	AF070623	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.00	1.00
	445903	AI347487	Hs.132781	class I cytokine receptor	1.00	36.00
	445932	BE046441	Hs.333555	Homo sapiens clone 24859 mRNA sequence	2.41	2.88
	445982	BE410233	Hs.13501	pescadillo (zebrafish) homolog 1, contai	1.60	1.35
10	446078	AI339982	Hs.156061	ESTs	1.00	42.00
	446102	AW168067	Hs.317694	ESTs	1.00	1.00
	446157	BE270828	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	1.70	1.53
	446269	AW263155	Hs.14559	hypothetical protein FLJ10540	73.01	48.00
	446292	AF081497	Hs.279682	Rh type C glycoprotein	1.55	1.26
15	446293	AM20213	Hs.149722	ESTs	1.00	2.00
	446423	AW139555	Hs.150120	ESTs	1.10	4.19
	446428	AW082270	Hs.12496	ESTs, Weakly similar to ALU4_HUMAN ALU S	0.53	3.26
	446432	AI377320	Hs.150058	ESTs	1.00	5.00
	446528	AU076640	Hs.15243	nucleolar protein 1 (120kD)	1.36	1.31
20	446574	AI310135	Hs.335933	ESTs	3.89	72.00
	446519	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	32.03	20.23
	446636	AC002563	Hs.15767	citron (rho-interacting, serine/threonin	4.19	5.07
	446783	AW138343	Hs.141867	ESTs	2.82	9.47
	446839	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	110.28	28.00
25	446849	AU076617	Hs.16251	cleavage and polyadenylation specific fa	3.26	2.94
	446856	AI814373	Hs.164175	ESTs	6.38	11.30
	446872	X97058	Hs.16362	pyrimidinergic receptor P2Y, G-protein c	1.98	2.03
	446880	AI811807	Hs.108646	Homo sapiens cDNA FLJ14934 fis, clone PL	94.90	113.00
	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	1.67	3.90
	446989	AK001898	Hs.16740	hypothetical protein FLJ11036	2.82	3.12
30	447022	AW291223	Hs.157573	ESTs	1.00	170.00
	447033	AI357412	Hs.157601	ESTs	7.15	107.00
	447078	AW885727	Hs.9914	ESTs	47.24	24.00
	447081	Y13896	Hs.17287	potassium inwardly-rectifying channel, s	0.12	17.88
35	447131	NM_004585	Hs.17466	retinoic acid receptor responder (tazaro	0.97	1.48
	447149	BE299857	Hs.326	TAR (HIV) RNA-binding protein 2	1.24	1.26
	447153	AA805202	Hs.315562	ESTs	1.00	54.00
	447164	AF026941	Hs.17518	Homo sapiens cly5 mRNA, partial sequence	1.00	67.00
	447178	AW594641	Hs.192417	ESTs	3.42	50.00
40	447250	AI878909	Hs.17883	protein phosphatase 1G (formerly 2C), ma	1.60	1.52
	447289	AW247017	Hs.36978	melanoma antigen, family A, 3	1.00	1.00
	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m	146.62	51.00
	447350	AI375572	Hs.172634	ESTs	1.00	12.00
45	447377	N27687	Hs.334334	transcription factor AP-2 alpha (activat	2.55	63.00
	447415	AW837335	Hs.28149	ESTs, Weakly similar to KF38_HUMAN KINES	0.91	1.13
	447425	AI963747	Hs.18573	acylphosphatase 1, erythrocyte (common)	1.00	35.00
	447519	U46258	Hs.339665	ESTs	59.89	49.00
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
50	447534	AA401369	Hs.190721	ESTs	1.00	17.00
	447636	Y10043		high-mobility group (nonhistone chromoso	1.41	1.11
	447688	N87079	Hs.19236	Target CAT	1.00	39.00
	447733	AF157482	Hs.19400	MAD2 (mitotic arrest deficient, yeast, h	1.17	1.12
	447769	AW873704	Hs.320831	Homo sapiens cDNA FLJ14597 fis, clone NT	6.47	5.95
55	447802	AW593432	Hs.161455	ESTs	0.73	2.34
	447850	AB018298	Hs.19822	SEC24 (S. cerevisiae) related gene famil	86.45	116.00
	447924	AI817226	Hs.313413	ESTs, Weakly similar to T23110 hypotheti	1.00	1.00
	447973	AB011169	Hs.20141	similar to S. cerevisiae SSM4	3.50	4.27
60	448030	N30714	Hs.325960	membrane-spanning 4-domains, subfamily A	4.13	142.00
	448105	AI538613	Hs.298241	Transmembrane protease, serine 3	1.15	2.24
	448243	AW369771	Hs.52620	Integrin, beta 8	15.84	1.00
	448278	W07369	Hs.11782	ESTs	0.97	1.90
	448290	AK002107	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	1.00	1.00
	448296	BE622756	Hs.10949	Homo sapiens cDNA FLJ14162 fis, clone NT	2.42	2.17
65	448357	BE274396	Hs.108923	RAB38, member RAS oncogene family	1.44	1.08
	448390	AL035414	Hs.21068	hypothetical protein	1.00	43.00
	448469	AW504732	Hs.21275	hypothetical protein FLJ11011	2.63	2.49
	448569	BE382657	Hs.21486	signal transducer and activator of trans	1.84	2.53
	448663	BE614599	Hs.106823	hypothetical protein MGC14797	3.29	46.00
70	448672	AI955511	Hs.225106	ESTs	1.00	21.00
	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	1.82	1.08
	448741	BE614567	Hs.19574	hypothetical protein MGC5469	2.48	1.92
	448757	AI366784	Hs.48820	TATA box binding protein (TBP)-associate	23.53	20.00
	448775	AB025237	Hs.388	nudix (nucleoside diphosphate linked mol	2.34	1.97
75	448826	AI580252	Hs.293246	ESTs, Weakly similar to putative p150 [H	74.07	62.67
	448830	AL031658	Hs.22181	hypothetical protein dJ310013.3	1.37	1.31
	448844	AI581519	Hs.177164	ESTs	1.00	31.00
	448988	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.84	1.95
	448993	AI471630		KIAA0144 gene product	1.63	1.49
80	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
	449029	N28989	Hs.22891	solute carrier family 7 (cationic amino	1.97	2.26
	449040	AF040704	Hs.149443	putative tumor suppressor	0.97	1.56
	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	27.13	90.00
	449053	AI625777	Hs.344766	ESTs	8.33	44.00
	449054	AF148848	Hs.22934	myoneurin	73.85	104.00
85	449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00

	449167	T05095	Hs.19597	KIAA1694 protein	1.61	2.36
	449207	AL044222	Hs.23255	nucleoporin 155KD	2.36	1.56
	449228	AJ403107	Hs.148590	protein related with psoriasis	1.15	1.15
5	449230	BE613348	Hs.211579	melanoma cell adhesion molecule	206.65	151.00
	449305	AI638293		gbt109b07.x1 NCL_CGAP_GC6 Homo sapiens	17.28	45.00
	449318	AW236021	Hs.78531	Homo sapiens, Similar to RIKEN cDNA 5730	26.39	35.00
	449448	D60730	Hs.57471	ESTs	1.00	1.00
	449467	AW205006	Hs.197042	ESTs	1.00	1.00
10	449523	NM_000579	Hs.54443	chemokine (C-C motif) receptor 5	56.80	216.86
	449722	BE280074	Hs.23960	cyclin B1	150.03	1.00
	449976	H08350	Hs.135056	Human DNA sequence from clone RP5-850E9	2.16	2.85
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitter)	1.17	1.45
	450098	W27249	Hs.8109	hypothetical protein FLJ21080	1.79	2.38
15	450101	AV649989	Hs.24385	Human hbc647 mRNA sequence	1.00	69.00
	450149	AW696781	Hs.132863	Zic family member 2 (odd-paired Drosophila)	1.00	1.00
	450193	AJ916071	Hs.15607	Homo sapiens Fanconi anemia complemental	29.85	34.00
	450221	AA328102	Hs.24641	cytoskeleton associated protein 2	1.00	1.00
	450372	BE218107	Hs.202436	ESTs	1.00	1.00
20	450375	AA039647	Hs.8850	a disintegrin and metalloproteinase domain	51.26	93.00
	450447	AF212223	Hs.25010	hypothetical protein P15-2	123.20	181.00
	450568	AL050078	Hs.25159	Homo sapiens cDNA FLJ10784 fis, clone NT	1.00	19.00
	450589	AI701505	Hs.202526	ESTs	1.00	23.00
	450684	AA872605	Hs.25333	Interleukin 1 receptor, type II	1.00	100.00
25	450701	H39960	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.89	1.55
	450705	U90304	Hs.25351	Iroquois homeobox protein 2A (IRX-2A) (1.00	45.00
	450832	AA401359	Hs.190721	ESTs	25.17	17.00
	450937	R49131	Hs.26267	ATP-dependant interferon response protein	90.92	90.00
	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	3.33	1.70
30	451105	AI761324		gbw160b11.x1 NCL_CGAP_Co16 Homo sapiens	15.02	124.00
	451110	AI955040	Hs.265398	ESTs, Weakly similar to transformation-	1.00	143.00
	451253	H48299	Hs.26126	claudin 10	3.02	2.29
	451291	R39288	Hs.6702	ESTs	1.00	1.00
	451320	AW498974		diacylglycerol kinase, zeta (104kD)	2.92	18.00
35	451380	H09280	Hs.13234	ESTs	6.90	6.67
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant)	35.75	72.00
	451437	H24143	Hs.31945	hypothetical protein FLJ11071	1.00	69.00
	451462	AK000367	Hs.26434	hypothetical protein FLJ20360	1.83	2.10
	451524	AK001466	Hs.26516	hypothetical protein FLJ10604	1.13	1.07
40	451541	BE279383	Hs.26557	plakophilin 3	1.88	1.33
	451592	AI805416	Hs.213897	ESTs	1.00	1.00
	451635	AA018899	Hs.127179	cryptic gene	1.52	1.92
	451743	AA401369	Hs.190721	ESTs	4.95	17.00
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	13.55	31.00
45	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	451871	AI821005	Hs.118599	ESTs	1.81	2.53
	451952	AL120173	Hs.301663	ESTs	1.00	22.00
	452012	AA307703	Hs.279766	kinesin family member 4A	3.43	2.26
	452046	AB018345	Hs.27657	KIAA0802 protein	56.59	19.00
50	452194	AI694413	Hs.332649	olfactory receptor, family 2, subfamily	1.67	4.09
	452206	AW340281	Hs.33074	Homo sapiens, clone IMAGE:3606519, mRNA,	9.31	53.00
	452240	AA401369	Hs.190721	ESTs	13.42	17.00
	452256	AK000933	Hs.28661	Homo sapiens cDNA FLJ10071 fis, clone HE	39.03	94.00
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	153.01	340.00
55	452291	AF015592	Hs.28853	CDC7 (cell division cycle 7, S. cerevisiae)	1.95	23.00
	452295	BE379936	Hs.28866	programmed cell death 10	42.33	61.00
	452304	AA025386	Hs.61311	ESTs, Weakly similar to S10590 cysteine	1.17	2.14
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodomain	1.00	13.00
	452349	AB028944	Hs.29189	ATPase, Class VI, type 11A	1.09	1.42
60	452367	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	54.49	53.00
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced protein	1.00	32.00
	452410	AL133619		Homo sapiens mRNA; cDNA DKFZp434E2321 (I	1.26	1.99
	452461	N78223	Hs.108106	transcription factor	24.47	35.00
	452571	W31518	Hs.34665	ESTs	54.61	102.00
65	452613	AA461599	Hs.23459	ESTs	1.39	1.32
	452699	AW295390	Hs.213062	ESTs	1.00	26.00
	452705	H49805	Hs.246005	ESTs	1.00	1.00
	452747	AF160477	Hs.61460	Ig superfamily receptor LNIR	112.87	1.29
	452787	AW294022	Hs.222707	KIAA1718 protein	1.00	1.00
70	452795	AW392555	Hs.18878	hypothetical protein FLJ21620	1.00	1.00
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	7.91	75.00
	452833	BE559681	Hs.30736	KIAA0124 protein	3.16	1.92
	452838	U65011	Hs.30743	preferentially expressed antigen in melanoma	174.35	1.00
	452862	AA401369	Hs.190721	ESTs	98.26	17.00
75	452865	AW173720	Hs.345805	ESTs, Weakly similar to A47582 B-cell growth	1.55	1.00
	452934	AA581322	Hs.4213	hypothetical protein MGC16207	1.73	1.19
	452946	X95425	Hs.31092	EphA5	1.00	1.00
	452976	R44214	Hs.101189	ESTs	1.58	1.98
	453028	AB006532	Hs.31442	RacQ protein-like 4	1.80	1.60
80	453095	AW295660	Hs.252766	ESTs	0.77	1.50
	453102	NM_007197	Hs.31664	Itz1 (Drosophila) homolog 10	1.00	1.00
	453103	AJ301052	Hs.153444	ESTs	1.00	1.00
	453120	AA292891	Hs.31773	pregnancy-induced growth inhibitor	1.23	1.20
	453153	N53893	Hs.24360	ESTs	1.00	83.00
85	453160	AI263307	Hs.239884	H2B histone family, member L	1.00	30.00
	453197	AB916269	Hs.109057	ESTs, Weakly similar to ALU5_HUMAN ALU S	1.00	134.00

5	453210	AL133161	Hs.32360	hypothetical protein FLJ10867	1.69	1.93		
	453240	AI989564	Hs.166254	hypothetical protein DKFZp566i133	1.00	1.00		
	453317	NM_002277	Hs.41696	keratin, hair, acidic, 1	1.19	1.27		
	453323	AF034102	Hs.32951	solute carrier family 29 (nucleoside tra	4.90	4.11		
	453331	AJ240665	Hs.8850	ESTs	199.42	340.00		
	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00		
	453431	AF094754	Hs.32973	glycine receptor, beta	1.00	1.00		
	453439	AI572438	Hs.32976	guanine nucleotide binding protein 4	3.44	5.17		
	453459	BE047032	Hs.257789	ESTs	2.84	5.58		
10	453563	AW608906.comp	Hs.181163	hypothetical protein MGC5629	4.58	90.00		
	453633	AA357001	Hs.34045	hypothetical protein FLJ20764	1.74	1.60		
	453775	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	19.49	1.00		
	453830	AA534296	Hs.20953	ESTs	24.92	25.00		
	453857	AL080235	Hs.35861	DKFZP586E1621 protein	167.59	66.00		
15	453867	AI929383	Hs.33032	hypothetical protein DKFZp434N185	1.00	39.00		
	453883	AI638516	Hs.347524	cofactor required for Sp1 transcription	1.97	1.58		
	453884	AA355925	Hs.36232	KIAA0186 gene product	63.89	20.00		
	453900	AW003582	Hs.226414	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.41	16.00		
	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1	7.09	22.00		
20	453941	U39817	Hs.36820	Bloom syndrome	29.75	19.00		
	453964	AI961486	Hs.12744	ESTs	1.00	1.00		
	453968	AA847843	Hs.62711	Homo sapiens, clone IMAGE:3351295, mRNA	2.06	1.81		
	453976	BE463830	Hs.163714	ESTs	3.02	131.00		
25	454024	AA935327	Hs.293907	hypothetical protein FLJ23403	1.00	131.00		
	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member	1.23	1.02		
	454042	T19228	Hs.172572	hypothetical protein FLJ20093	30.63	171.00		
	454059	NM_003154	Hs.37048	stathmin	1.00	1.00		
	454066	X00358	Hs.37058	calcitonin/calcitonin-related polypeptide	1.01	1.45		
	454098	W27953	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11		
30	454241	BE144666	Hs.110826	gb:CM2-HT0176-041099-017-c02 HT0176 Homo	6.33	5.04		
	454417	AI244459	Hs.110826	trinucleotide repeat containing 9	4.30	7.82		
	454439	AW819152	Hs.154320	DKFZP566O1646 protein	1.00	1.00		
	455175	AW993247	Hs.154320	gb:RC2-BN0033-180200-014-h09 BN0033 Homo	13.75	103.00		
35	455601	AI368680	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00		
	456237	AA203682	Hs.816	gb:zx52e07.r1 Soares_fetal_liver_spleen_	1.00	1.00		
	456321	NM_001327	Hs.87225	cancer/testis antigen	1.14	1.10		
	456475	NM_000144	Hs.95938	Friedreich ataxia	1.00	48.00		
	456508	AA502764	Hs.123469	ESTs, Weakly similar to AF208855 1 BM-01	162.25	189.00		
40	456534	X91195	Hs.100623	phospholipase C, beta 3, neighbor pseudo	2.12	1.80		
	456736	AW248217	Hs.1619	achaete-scute complex (Drosophila) homolog	1.15	1.94		
	456759	BE259150	Hs.127792	delta (Drosophila)-like 3	1.00	1.00		
	456990	NM_004504	Hs.171545	HIV-1 Rev binding protein	16.42	84.00		
	457200	U33749	Hs.197764	thyroid transcription factor 1	0.57	1.76		
45	457234	AW968360	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	2.71	4.15		
	457465	AW301344	Hs.122908	DNA replication factor	46.37	47.00		
	457489	AI693815	Hs.127179	cryptic gene	1.12	1.35		
	457646	AA725650	Hs.112948	ESTs	1.55	2.51		
	457733	AW974812	Hs.291971	ESTs	1.00	55.00		
50	457819	AA067484	Hs.35406	ESTs, Highly similar to unnamed protein	4.36	3.18		
	458092	BE545684	Hs.343566	KIAA0251 protein	1.00	1.32		
	458098	BE550224	Hs.343566	metallothionein 1E (functional)	1.00	22.00		
	458207	T28472	Hs.7655	U2 small nuclear ribonucleoprotein and	2.06	1.88		
	458242	BE299588	Hs.28465	Homo sapiens cDNA: FLJ21869 fis, clone H	1.00	1.00		
55	458247	R14439	Hs.209184	ESTs	7.00	9.85		
	458679	AW975460	Hs.142913	ESTs	1.00	3.00		
	458776	AW451034	Hs.326525	arylsulfatase D	1.31	2.01		
	458933	AI638429	Hs.24783	RAN binding protein 1	1.98	1.71		
	459352	AW810363	Hs.206828	ESTs	12.60	63.00		
	459670	F01020	Hs.172004	titin	1.00	1.00		
60	459702	AI204895	Hs.172004	gb:an03c03.x1 Stralagene schizo brain S1	1.00	237.00		

TABLE 9B

65	Pkey:	Unique Eos probeset identifier number
	CAT number:	Gene cluster number
	Accession:	Genbank accession numbers
70	Pkey	CAT Number
	407746	10125_1
		AK001962 R69415 BE464605 AA418699 AA053293 AA149075 AA058396 AW338226 AW272659 AA454607 AI139535 AW469852 AI275461
		AW271982 AA730033 AA576507 AA991217 AA782067 AI985851 AA805864 AA505598 AW469857 R69546 AA988279 AW001647 N63320
		D82661 T27343 AA306950 AA360989 R58778
	408070	1036888_1
	408660	107294_1
75	409522	113735_1
	409866	1156522_1
	410032	1170435_1
	411089	123172_1
		AW148852 BE350895
		AA525775 AA056342 AI538978 AW975281 AA664986
		AA075382 AA075431
		AW502152 H14202 H29772
		BE065985 BE065944 BE066008 BE066083 BE066093
		AA456454 AA713730 AA091294 AA584921 N86077 AW836781 AA601031 AA579876 AA551106 AA633188 AW905577 AI955808 AI679386
		AI679895 AA514764 AA454562 AI082362 AA595822 AA551351 AA586369 AA666384 AA188934 AA666398 AA551297 AA565188
80	411152	1234028_1
		BE069199 AW936012 AW877466 AW819782 AW935798 AW835546 AW936042 BE069121 AW835625 AW877536 AW935885 BE069202
		AW820019 AW935937 BE160180 AW935946 BE069101 BE069125 AW877527 BE160316 BE160393 AW935794 AW835701 AW935784
		AL031778 X59711 NM_002505 M59079 AI870439 AI494259 AW664010 AA405063 AA436132 BE174516 AA412691 AI400314 AA436024
		T29403 BE079412 BE079428 N90322 AI631202 AI141758 AI016793 AI167566 AI862075 AI375230 AI208445 AW235763 AL044113 AA382556
		AW953918 AA927051 AA889823 BE003094 AW390155 AW360805 AW360823 AW360810 AA425472 AI694282 AL044114 AI684577 AI809865

			AI478773 AI160445 AI674630 N69088 AW665529 N49278 AI129239 AI457890 AI621264 AW297152 AI268215 AA907787 AI266170 AI017982 AI963541 AI469807 AI969353 BE552356 N66509 AA736741 AA382555 AW075811 AW292026 H06382 AW957730 AA352014 R13591 AA121201 D60420 BE263253 BE047862 Z41952 AI424991 AI693507 AI853108 AA599060 AI091148 AA598689 R39887 AA813482 AW016452 H06383 R41807 AI364268 AA620528 AI241940 AW089149 AW090733 AW088875 Z38240 AA121202 R17734 BE157489 BE157560
5	412811	132943_1	AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245 AA082436 H72525 H77575 NA9786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW960304 T29812 AA476873 BE297387 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 AI719075 BE270172 BE269819 AA889955 AI204630 W25243 AI935150 AA872039 W72395 T99630 AI422691 H98460 N31428 BE255916 H03265 AI857576 AA776920 AA910644 AA459522 AA293140 AW514667 R75953 AW662396 AA662522 AI865147 AI423153 AW262230 AA584410 AA583187 AW024595 AW069734 AI828996 AA282997 AA876046 AW613002 AA527373 AW972459 AI831360 AA621337 AA100926 AA772418 AA594628 AI033892 W95096 AI034317 AA398727 AI085031 N95210 AI459432 AI041437 AA932124 AA627684 AA935829 AI004827 AI423513 AI094597 H42079 R54703 AI630359 AA617681 AA978045 AA643280 W44561 AI991988 AI537692 AI090282 AA740817 AI312104 AI911822 AA416871 AI185409 AA129784 AA701623 AI075239 AI139549 AA633648 AI339996 AI336880 AA399239 AI078708 AI085351 AI362835 AI346618 AI146955 AI989380 AI348243 N92892 AA6765850 AI494230 AI278887 AA962596 AI492600 W80435 AA001979 R97424 AI129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785 AI494211 AW059601 AW886710 R92790 N59755 AI361128 AW589407 H47725 H97534 H48076 H48450 N96317 AA9300758 H03431 R76789 AA954344 H77576 R96823 AI457100 N92845 NA9682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923 H03266 BE261919 AA769633 AA480310 AA507454 AA910586 AI203723 AW104725 W25611 W25071 T88980 H03513 T77589 R99156 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283872
10	413690	1383256_1	AI267700 AI720344 AA191424 AI023543 AI469833 AA172056 AW958465 AA172236 AW953397 AA355086
15	414883	15024_1	AW265494 AA455904 AA195677 AW265432 AW991605 AA456370 N28754 N28747 AI568146 AI979339 AA322671 AA322672 AW955043 AI990326 AA776406 AI016250 AAB43678 AW451882 N23137 N23129 W70051 AI038748 AAB31327 AI925845 AW945895 Z42183 T31621 T97478 D62703 AA242966 D79798
20	415989	156454_1	AU076704 T74854 T74860 T72098 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72826 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M54982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T58035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74517 T71958 T69440 T61875 R06796 H48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AAQ11465 AA345378 AV654847 AV654272 AV656001 AI064740 T82897 N33594 AA344542 AW805054 AI207457 T61743 AA026737 H94389 AA382695 AA918409 T68044 S82092 T39959 AI017721 AA312395 AA312919 T40156 H66239 AV652989 H38728 R98521 AV655200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343596 AW470774 AV651256 N54417 AA812862 AW182929 AI111192 H61463 H72060 AA344503 H38639 AI277511 AV661108 AI207625 T47810 AA235252 T27853 T47778 R95746 H70820 AA701463 AW827166 R98475 C20925 AV657287 T71959 T71313 T73920 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T65778 AA344726 T72854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 AI110639 AA344603 AF063513 T64696 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74677 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA342807 AA341964 T53747 T72042 T62764 AI064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 AI207557 T68219 W86031 T69081 T64232 R93196 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 AI248502 R29454 T64764 T57001 T73052 T71429 T51176 T58886 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 AI792788 BE142230 AA252019
25	417324	166714_1	AI910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 AI571948 AA507595 AA614579 AA587613 R83818 AA586312 AA614409 AA307578 AI925552 AW950155 AI910083 M12075 BE074052 AW004668 AA578674 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA858276 R55389 AI001051 AW050700 AW750216 AA614539 BE074045 AI307407 AW602303 BE073575 AI202532 AA524242 AI970839 AI909751 BE076078 AI909749 R55292 AW881145 AA490718 M85637 AA304575 T06067 AA331991 AL119930 AA320696 AW752565 AL031985 AL137241 AT792386 AT733664 AI857654 AI049911 AA337221 AA336756 AW966196 AW953120 R56325 AA349562 AI493134 AI498691 AW771508 AI498457 AI768408 AI783624 AI383985 AI580267 D79813 AA393768 AK001536 AA191092 AW510354 AI554256 AL353968 AA134266 AA663848 AA400100 AA401424
30	418712	1784125_1	AL038843 AA161338 BE268213 AA425597 N87306 AA092969 BE566038 AA247451 N47392 AI928802 AW182584 AW027872 AI819831 AI936994 W56258 AI653448 AI278611 AI283557 AI824306 AW339658 AW150899 AA687514 N47393 N29885 AA973459 AI038904 AI292064 AI034339 AW674593 N72156 AI079733 AI038683 AI291616 AA491599 AA933675 AA837380 BE006554 BE006473 AI087090 T33044 AA652043 AI203503 AA583959 W35283 AI129926 Z41844 AW020925 AW575848 AI684603 AA493297 AI140589 AI277175 AA425444 AI932767 W02632 BE396786 R37261 AW207206 AW341473 AA448195 AI951341 AA249027 AL038984 AK001993 AL080066 AV652725 BE566226 AA345557 AA315222 AA090585 AA375688 AA301092 AA298454 W05762 AW607939 H51658 D83880 N84323 BE296821 AW947007 D81461 AW079261 AA329482 AW901780 AI354442 AA772275 R31663 AI354441 AI767525 H92431 AI916735 H93575 AI394255 AW014741 AI573090 C06195 AW612857 AW265195 AI339558 AI377532 AI038821 AI919424 AI589705 AW055215 AI336532 AI338051 AA806547 C75509 C00618 AW071172 AW769904 AA630381 AI678018 AI863985 T79662 BE221049 AW265018 AI589700 AW196655 H76573 AI370908 BE042393 N75017 AI698870 AW960115 AL133561 AL041090 AL117481 AL122069 AW439292 AI958826 AW072916 AI184913 AA489195 AW466994 AW469044 N59350 AI819642 AI280239 AI220572 AA789302 AI473611 AW841126 D60937 BE041395 AA491826 AA621946 AA715980 AA666102 AW970622 AA503009 AA502998 AA502989 AA502805 T92188 AA221036 R87170 BE537068 BE544757 C18935 AW812058 T92565 AA227415 AA233942 AA223237 AA668403 AA601627 AW866639 BE061833 BE000620 AW961170 AW847519 AA308542 AW821833 AW945688 C04699 AA205504 AA377241 AW821667 AA055720 AW817981 AW856468 AA155719 AA179928 T03007 AW754298 AA227407 AA113928 AA307904 C16859 AI798376 S46400 AW811617 AW811616 W00557 BE142245 AW858232 AW861851 AW858362 AA232351 AA218567 AA055556 AW858231 AW857541 AW814172 H66214 AW814398 AF134164 AA243093 AA177345 AA199942 AA223384 AA227092 AA227080 T12379 AA092174 T61139 AA149776 AA698829 AW879188 AW813567 AW813538 AI267168 AA157718 AA157719 AA100472 AA100774 AA130756 AA157705 AA157730 AA157715 AA053524 AW849581 AW854566 C05254 AW882836 T92637 AW812621 AA206583 AA209204 BE156909 AA226824 AI829309 AW991957 N66951 AA527374 H66215 AA045564 AI694265 H60808 AA149726 AW195620 BE081333 BE073424 AW817662 AW817705 AW817703 AW817659 BE081531 H59570 AA626980 AI126603 BE504035
35	419443	184788_1	
40	419502	18535_1	
45			
50	419936	189181_1	
55	421582	2041_1	
60	422128	211994_1	
65	423034	224122_1	
70	423816	23234_1	
75	424200	236595_1	
80	424999	245835_1	
85	426966	273896_1	
	426991	27415_1	
	427260	276598_1	
	428023	28589_2	

438091	44954_1	AW373062 T55662 AI299190 BE174210 AW579001 H01811 W40186 R67100 AI923886 AW952164 AA628440 AW898607 AW898616 AA709126 AW898628 AW898644 AA947932 AW898625 AW898622 AI276125 AI185720 AW510698 AA987230 T52522 BE467708 AW243400 AW043642 AI288245 AI186932 D52654 D55017 D52715 D52477 D53933 D54679 AI298739 AI146984 AI922204 N98343 BE174213 AA845571 AI813854 AI214518 AI635262 AI139455 AI707807 AI698085 AW884528 AI024768 AI004723 AW087420 AI565133 N94964 AI268939 AW513280 AI061126 AI435818 AI859106 AI360506 AI024767 AA513019 AA757598 X56196 AA902959 AI334784 AI860794 AA010207 AW890091 AW513771 AI951391 AI337671 T52499 AA890205 AI640908 H75966 AA463487 AA358688 AI961767 AI865295 AA780994 AI985913 BE174196 AA029094 AW592159 T55581 N79072 AI611201 AA910812 AI220713 AW149306 AI758412 AA045713 R79750 N76096 AW979121 AA847986 AA28098 AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882 AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077 AL109688 R23665 R26578 AA570256 AW014761 AA573721 AJ473237 AI022165 AA554071 AA127551 N90525 AW973623 AA447991 AA243852 BE328850 AI148171 AI359627 AI005068 AI356567 AA232991 AW016855 AA906902 AA233101 AA127550 BE512923 AI188710 AI032142 AW078833 N30308 AW675632 AI219028 AI341201 N22181 H95390 W24187 W24194 R17789 Y10043 NM_005342 L05085 AL034450 BE614226 AW749053 AA379173 AA248230 BE514634 AA334622 R70656 AA367593 AA214649 AA369318 AW957081 R05760 AA039903 AI886597 AW630122 AA906264 AA041527 R01145 AI086688 BE463637 AA398795 AI354883 AI768938 AI589996 AI452952 AI168582 AI189869 AI086670 AW262560 AW613854 AA862839 AA435840 AA670197 AI024032 AI990659 AI990089 N81095 AA847919 AW960150 AA211075 AA044704 AA367594 AW582587 AW858854 AW818630 AW818281 AW818433 AW582595 AA096002 N83992 AI471630 BE540637 BE265481 AW407710 BE513882 BE546739 AA053597 BE140503 BE218514 AW956702 AI656234 AI636283 AI567265 AW340858 BE207794 AA053085 R69173 AA292343 AA454908 AA293504 AI659741 AI927478 AA399460 AI760441 AA346416 BE047245 AA730380 AA394063 AA454833 AI982791 AI557270 AI813332 AI767858 AA427705 D20284 AI221458 BE048537 AI263048 AA346417 AA911497 BE537702 AI638293 AW813561 AI761324 AW880941 AW880937 AW118072 AI631982 T15734 AA224195 AI701458 W20198 F26326 AA890570 N90552 AW071907 AI671352 AI375892 T03517 R88265 AI124088 AA224388 AI084316 AI354686 T33652 AI140719 AI720211 T03490 AI372637 T15415 AW205836 AA630384 T03515 T33230 AA017131 AA443303 T33623 AI222556 T33511 T33785 AI419606 D55612 W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW450652 AW449519 AA993634 AI806539 AA351618 AW449522 AI827626 AA904788 AA380381 AA886045 AA774409 BE003229 Z41756 AL133619 AA468118 AA383064 AI476447 T09430 AI673758 AA524895 AI581345 AI300820 AW498812 AA256162 AI559724 AI685732 AA602400 AA905453 AI204595 AW166541 AA157456 AA156269 AA383652 AA431072 AW592707 AI435410 AW272464 AI215594 AA622747 R74039 N35031 AI804128 AW513621 AA668351 AI026826 AI493388 AA614641 W81604 AI567080 AI214351 AA730140 AI125754 AI200813 AI269603 AI565082 AI807095 AI476629 AA505909 AI368449 AI686077 AI582930 AW085038 AA757863 AA730154 AI767072 AA468316 AI734130 AI734138 AA426284 AA433997 AI741241 AW043563 AI732741 AI732734 AA437369 AA425820 AA664048 R74130 BE144668 BE184942 AW238414 BE184946 AW993247 AW861464 AA203682 R11958 BE550224 AA832519 N45402 AW885857 N29245 BE465409 W07677 AW970089 AI299731 AA482971 BE503548 H18151 W79223 AF086393 AA461301 W74510 R34182 AI090689 N46003 BE071550 R28075 AW134982 AI240204 AI138906 AW026179 AI572316 BE466182 AI206395 AI276154 AI273269 AI422817 AI371014 AI421274 AI188525 AA939164 BE549810 AW137865 AI694996 BE503841 AA459718 BE327407 BE467534 BE218421 BE467767 AA989054 BE467053 AI797130 BE327781
--------	---------	--

TABLE 9C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

55	Pkey	Ref	Strand	Nt_position
	400512	9798593	Minus	1439-1615
	400517	9796686	Minus	4996-50346
	400560	9843598	Plus	94182-94323,97056-97243,101095-101238,102824-103005
	400664	8118496	Plus	13558-13721,13942-14090,14554-14679
	400665	8118496	Plus	16879-17023
60	400666	8118496	Plus	17982-18115,20297-20456
	400749	7331445	Minus	9162-9293
	400763	8131616	Minus	35537-35784
	401027	7230983	Minus	70407-70554,71060-71160
	401093	8516137	Minus	22335-23166
65	401203	9743387	Minus	172961-173056,173868-173928
	401212	9858408	Plus	87839-88028
	401411	7799787	Minus	144144-144329
	401435	8217934	Minus	54508-55233
	401464	6682291	Minus	170689-170834
70	401714	6715702	Plus	96484-96681
	401747	9789672	Minus	118596-118816,119119-119244,119609-119761,120422-120990,130161-130381,130468-130593,131097-131258,131866-131932,132451-132575,133580-134011
	401760	9929699	Plus	83126-83250,85320-85540,94719-95287
	401780	7249190	Minus	28397-28617,28920-29045,29135-29296,29411-29567,29705-29787,30224-30573
75	401781	7249190	Minus	83215-83435,83531-83655,83740-83901,84237-84393,84955-85037,86290-86814
	401785	7249190	Minus	165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
	401797	6730720	Plus	6973-7118
	401961	4581193	Minus	124054-124209
	401985	2580474	Plus	61542-61750
80	401994	4153858	Minus	42904-43124,43211-43336,44607-44763,45199-45281,46337-46732
	402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
	402260	3399665	Minus	113765-113910,116653-116765,116808-116940
	402265	3287673	Plus	21059-21168
	402297	6599824	Plus	35279-35405,35573-35659
85	402408	9796239	Minus	110326-110491

	402420	9796339	Plus	129750-129919
	402674	8077108	Minus	39290-39502
	402802	3287156	Minus	53242-53432
5	402994	2996643	Minus	4727-4969
	403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
	403306	8099945	Plus	127100-127251
	403329	8516120	Plus	96450-96598
	403381	9438267	Minus	26009-26178
10	403478	9958258	Plus	116458-116554
	403485	9966528	Plus	2888-3001,3198-3532,3655-4117
	403627	8569879	Minus	23868-24342
	403715	7239669	Plus	85128-85292
	404044	9558573	Minus	225757-225939
	404076	9931752	Minus	3848-3967
15	404101	8076925	Minus	125742-125997
	404140	9843520	Plus	37761-38147
	404185	9926489	Minus	69025-69128
	404185	4572584	Minus	129171-129327
20	404210	5006246	Plus	169926-170121
	404253	9367202	Minus	55675-56055
	404287	2326514	Plus	53134-53281
	404298	9944263	Minus	73591-73723
	404347	9838195	Plus	74493-74829
25	404440	7528051	Plus	80430-81581
	404721	9856648	Minus	173763-174294
	404784	4826439	Plus	101619-101898
	404854	7143420	Plus	14260-14537
	404877	1519284	Plus	1095-2107
30	404927	7342002	Plus	68690-69563
	404998	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40674,42351-42450
	405449	7622497	Plus	42236-42570
	405568	6006906	Plus	35912-36065
	405572	3800891	Plus	85230-85938
35	405646	4914350	Plus	741-969
	405676	4557087	Plus	73195-73917
	405770	2735037	Plus	61057-62075
	405832	7767812	Minus	123525-123713
	406137	9166422	Minus	30487-31058
40	406360	9256107	Minus	7513-7673
	406399	9258288	Minus	63448-63554
	406467	9795551	Plus	182212-182958

45 TABLE 10A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer and Non-malignant Lung Disease
Table 2A shows about 307 genes up-regulated in non-malignant lung disease relative to lung tumors and normal body tissues and/or down-regulated in lung tumors relative to normal lung and non-malignant lung disease. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

50 Table 10B show the accession numbers for those Pkey's lacking UnigenelD's for table 10A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

55 Table 10C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 10A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

60 Pkey_i Unique Eos probeset identifier number
ExAccn: Exemplar Accession number, Genbank accession number
UnigenelD: UnigenelD number
Unigenel Title: Unigenel gene title
R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

	Pkey	ExAccn	UnigenelD	Unigenel Title	R1	R2
65	404394			ENSP00000241075:TRRAP PROTEIN.	0.79	3.10
	404916			Target Exon	1.00	159.00
	405257			Target Exon	1.00	422.00
70	407228	M25079	Hs.155376	hemoglobin, beta	0.47	2.33
	407568	AA740964	Hs.62699	ESTs	1.00	123.00
	408562	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein.	1.00	230.00
	409031	AA376836	Hs.76728	ESTs	1.00	128.00
	410434	AF051152	Hs.63668	tol-like receptor 2	39.65	149.00
75	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	1.00	109.00
	410808	T40326	Hs.167793	ESTs	1.14	13.14
	412351	AL135960	Hs.73828	T-cell acute lymphocytic leukemia 1	0.37	2.27
	412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
	413795	AL040178	Hs.142003	ESTs	0.10	11.90
80	414154	AW205314	Hs.323060	ESTs	0.62	2.09
	414214	D49958	Hs.75819	glycoprotein MGA	0.03	4.55
	414998	NM_002543	Hs.77729	oxidised low density lipoprotein (lecitin	0.64	2.97
	415122	D60708	Hs.22245	ESTs	0.07	8.97
	415765	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and	0.67	1.65
85	415775	H00747	Hs.29792	ESTs, Weakly similar to I38022 hypothetical	0.29	2.64
	415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00

	416319	AI815601	Hs.79197	CD83 antigen (activated B lymphocytes, i	15.32	237.00
	416402	NM_000715	Hs.1012	complement component 4-binding protein,	0.64	4.00
	417355	D13168	Hs.82002	endothelin receptor type B	0.01	3.90
5	417421	AI138201	Hs.82120	nuclear receptor subfamily 4, group A, m	36.30	357.00
	417511	AI049176	Hs.82223	chordin-like	1.00	179.00
	418489	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h	0.02	6.00
	418726	BE241812	Hs.87860	protein tyrosine phosphatase, non-recept	1.00	113.00
	418741	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	0.44	1.90
	418883	BE387036	Hs.1211	acid phosphatase 5, tartrate resistant	0.96	2.04
10	419086	NM_000216	Hs.89591	Kallmann syndrome 1 sequence	0.62	2.74
	419150	T29618	Hs.89540	TEK tyrosine kinase, endothelial (venous	0.03	6.90
	419235	AW470411	Hs.288433	neurotrophin	1.48	5.13
	419407	AW410377	Hs.41502	hypothetical protein FLJ21276	37.55	336.00
	420556	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	0.80	3.65
15	420658	AA279098	Hs.187636	ESTs	1.65	8.07
	420729	AW964897	Hs.290825	ESTs	2.99	25.82
	421177	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	0.46	1.95
	422060	R20893	Hs.325823	ESTs, Moderately similar to ALL5_HUMAN A	1.00	156.00
	422426	W79117	Hs.58559	ESTs	0.03	7.44
20	422652	AW967969	Hs.118958	syntactin 11	0.14	3.62
	423099	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t	0.01	3.16
	424433	H04607	Hs.9218	ESTs	0.75	141.75
	424585	AA464840	Hs.131987	ESTs	1.00	167.00
25	424711	NM_005795	Hs.152175	calcitonin receptor-like	0.43	3.01
	424973	X92521	Hs.154057	matrix metalloproteinase 19	0.37	19.45
	425023	AW956889	Hs.154210	endothelial differentiation, sphingolipi	0.14	3.35
	425664	AJ006276	Hs.159003	transient receptor potential channel 6	1.00	94.00
	425998	AJ076629	Hs.165950	fibroblast growth factor receptor 4	0.68	1.42
30	426657	NM_015865	Hs.171731	solute carrier family 14 (urea transport	0.03	3.74
	426753	T89832	Hs.170278	ESTs	1.00	141.00
	427558	D49493	Hs.2171	growth differentiation factor 10	1.00	117.00
	427983	MI1706	Hs.2233	colony stimulating factor 3 (granulocyte	0.75	2.20
	428467	AK002121	Hs.184465	hypothetical protein FLJ11259	0.76	2.25
35	428927	AA441837	Hs.90250	ESTs	0.01	3.62
	429496	AA453800	Hs.192793	ESTs	1.00	138.00
	430468	NM_004673	Hs.241519	angiotensin-like 1	1.00	132.00
	431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	167.00
	431728	NM_007351	Hs.268107	multimerin	1.00	157.00
40	431848	AI378857	Hs.126758	ESTs, Highly similar to AF175283 1 zinc	0.34	2.24
	432128	AA127221	Hs.117037	ESTs	0.00	1.15
	432519	AI221311	Hs.130704	ESTs, Weakly similar to BCHUIA S-100 pro	0.01	2.06
	433043	W57554	Hs.125019	lymphoid nuclear protein (LAF-4) mRNA	1.00	267.00
	433803	AI823593	Hs.27688	ESTs	1.00	105.00
45	434730	AA644669	Hs.193042	ESTs	1.05	3.15
	435472	AW972330	Hs.283022	triggering receptor expressed on myeloid	0.83	1.94
	436532	AA721522		gbmv54h12.r1 NCI_CGAP_Ew1 Homo sapiens	1.00	218.00
	437119	AI379921	Hs.177043	ESTs	1.00	133.00
	437140	AA312799	Hs.283689	activator of CREM in testis	0.67	122.67
50	437211	AA382207	Hs.5509	ecotropic viral integration site 2B	1.00	142.00
	437960	AI669586	Hs.222194	ESTs	1.00	147.00
	438202	AW169287	Hs.22588	ESTs	1.00	141.00
	438873	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	0.71	3.66
	438875	AA827640	Hs.189059	ESTs	23.32	370.00
55	441048	AA913488	Hs.192102	ESTs	0.77	8.50
	441188	AW292830	Hs.255609	ESTs	3.43	16.36
	441499	AW298235	Hs.101689	ESTs	1.00	167.00
	444513	AI120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
	444527	NM_005408	Hs.11383	small inducible cytokine subfamily A (Cy	46.47	153.00
60	444561	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	0.01	3.08
	445279	R41900	Hs.22245	ESTs	0.60	141.00
	446017	N98238	Hs.55185	ESTs	0.18	2.39
	446584	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15	0.10	2.16
	446598	N99013	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.01	2.53
65	447367	AI375922	Hs.159367	ESTs	0.46	2.64
	448106	AI800470	Hs.171941	ESTs	18.05	296.00
	448253	H25899	Hs.201591	ESTs	1.00	141.00
	449275	AW450848	Hs.205457	perlecan	0.56	1.38
	450400	AI694722	Hs.279744	ESTs	0.88	4.33
70	450695	AI654223	Hs.16026	hypothetical protein FLJ23191	0.52	2.08
	450726	AW204600	Hs.250505	retinoic acid receptor, alpha	0.79	2.01
	451497	H83294	Hs.284122	Wnt inhibitory factor-1	0.35	2.03
	451533	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	0.13	2.25
	453636	R67837	Hs.169872	ESTs	1.00	116.00
75	458332	AI000341	Hs.220491	ESTs	1.00	192.00
	459580	AA022888	Hs.176065	ESTs	0.20	2.98
	400269			Eos Control	0.40	2.40
	403421			NM_016369*:Homo sapiens claudin 18 (CLDN	0.53	1.77
	407570	Z19002	Hs.37096	zinc finger protein 145 (Kruppel-like, e	0.01	3.18
80	412295	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1	0.56	1.74
	414517	M24461	Hs.76305	surfactant, pulmonary-associated protein	0.64	1.50
	417204	N81037	Hs.1074	surfactant, pulmonary-associated protein	0.33	1.16
	418307	U70867	Hs.83974	solute carrier family 21 (prostaglandin	0.53	1.55
	418935	T28499	Hs.89485	carbonic anhydrase IV	0.20	1.28
85	421502	AF111856	Hs.105039	solute carrier family 34 (sodium phospho	0.78	1.90
	421798	N74880	Hs.29877	N-acylsphingosine amidohydrolase (acid c	0.59	1.54

	423354	AB011130	Hs.127436	calcium channel, voltage-dependent, α	0.59	1.55
	423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	0.35	1.62
	425438	T62216	Hs.270840	ESTs	0.23	9.45
5	426828	NM_000020	Hs.172670	activin A receptor type II-like 1	0.03	1.71
	427019	AA001732	Hs.173233	hypothetical protein FLJ10970	0.01	1.49
	428043	T92248	Hs.2240	uteroglobin	0.42	1.26
	430280	AA361258	Hs.237868	Interleukin 7 receptor	0.46	2.43
	431433	X65018	Hs.253495	surfactant, pulmonary-associated protein	0.57	1.59
10	431723	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.29	1.80
	432985	T92363	Hs.178703	ESTs	0.32	2.27
	441835	AB036432	Hs.184	advanced glycosylation end product-speci	0.31	1.51
	442275	AW449467	Hs.54795	ESTs	0.55	1.78
	443709	AI082692	Hs.134662	ESTs	0.00	3.02
15	444325	AW152618	Hs.16757	ESTs	0.32	2.49
	450954	AI904740	Hs.25691	receptor (calcitonin) activity modifying	0.46	1.74
	451558	NM_001089	Hs.26630	ATP-binding cassette, sub-family A (ABC1	0.52	1.87
	453310	X70697	Hs.553	solute carrier family 6 (neurotransmitte	0.00	3.30
20	456855	AF035528	Hs.153863	MAD (mothers against decapentaplegic, Dr	0.01	2.31
	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane	0.66	2.20
	400754			Target Exon	1.00	297.00
	401045			C11001883*:g 6753278 ref NP_033938.1 c	1.00	109.00
	401083			NM_016582*:Homo sapiens peptide transpor	0.89	1.39
	402474			NM_004079:Homo sapiens cathepsin S (CTSS	1.45	4.47
25	402808			ENSP00000235229:SEMB.	1.00	1.87
	403021			C21000030:g 9955960 ref NP_063957.1 AT	1.00	149.00
	403438			NM_031419*:Homo sapiens molecule possess	1.06	2.96
	403687			NM_007037*:Homo sapiens a disintegrin-li	0.04	4.89
	403764			NM_005463:Homo sapiens heterogeneous nuc	1.00	225.00
30	404277			NM_019111*:Homo sapiens major histocompa	0.97	1.93
	404288			NM_002944*:Homo sapiens v-ros avian UR2	1.00	68.00
	404518	AI815601		CD83 antigen (activated B lymphocytes, l	0.02	1.83
	405106			C11001637*:g 5032241 ref NP_005732.1 z	1.00	235.00
	405381			Target Exon	1.00	93.00
35	406387			Target Exon	1.37	6.02
	406646	M33600		major histocompatibility complex, class	0.86	2.46
	406714	AI219304	Hs.266959	hemoglobin, gamma G	0.01	3.19
	406753	AA505665	Hs.217493	annexin A2	1.00	147.00
40	406973	M34996	Hs.198253	major histocompatibility complex, class	1.03	2.04
	407248	U82275	Hs.94498	leukocyte immunoglobulin-like receptor,	1.00	64.00
	407510	U96191		gb:Human trophoblast hypoxia-regulated f	1.00	90.00
	407731	NM_000066	Hs.38069	complement component 8, beta polypeptide	1.00	67.00
	407830	NM_001086	Hs.587	arylacetamide deacetylase (esterase)	1.00	102.00
45	408045	AW138959	Hs.245123	ESTs	1.00	70.00
	408074	R20723		ESTs	1.00	112.00
	408374	AW025430	Hs.155591	forkhead box F1	0.07	10.17
	409064	AA062954	Hs.141883	ESTs	0.39	2.31
	409083	AF060083	Hs.673	interleukin 12A (natural killer cell sti	1.00	95.00
50	409153	W03754	Hs.50813	hypothetical protein FLJ20022	0.01	4.55
	409203	AA780473	Hs.687	cytochrome P450, subfamily IVB, polypept	0.01	3.72
	409238	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp564G112 (fr	1.00	79.00
	409389	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	0.14	27.35
	409718	D86640	Hs.56045	src homology three (SH3) and cysteine ri	1.00	113.00
55	410798	BE178622	Hs.16291	gb:PM3-HT0605-270200-001-a02 HT0605 Homo	0.64	2.47
	411020	NM_006770	Hs.67726	macrophage receptor with collagenous str	0.55	2.40
	411667	BE160198		gb:QV1-HT0413-010200-059-h03 HT0413 Homo	1.00	111.00
	412000	AW576555	Hs.15780	ATP-binding cassette, sub-family A (ABC1	1.00	95.00
	412358	BE047490	Hs.24172	ESTs	1.00	87.00
60	412420	AL035668	Hs.73853	bone morphogenetic protein 2	1.43	8.07
	412564	X83703	Hs.31432	cardiac ankyrin repeat protein	0.02	3.07
	412869	AA290712	Hs.82407	CXC chemokine ligand 16	0.93	1.72
	412870	N22788	Hs.82407	CXC chemokine ligand 16	0.97	1.51
	413529	U11874	Hs.846	interleukin 8 receptor, beta	0.02	2.42
65	413533	BE145973		gb:QV4-HT0222-011199-019-e05 HT0222 Homo	0.65	1.50
	413689	BE157286	Hs.20631	zinc finger protein, subfamily 1A, 5 (Pa	20.87	232.00
	413724	AA131466	Hs.23767	hypothetical protein FLJ12666	1.00	80.00
	413800	AI129238	Hs.192235	ESTs	1.00	85.00
	413802	AW964490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-	1.00	213.00
70	413829	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)	0.02	3.93
	414376	BE393856	Hs.66915	ESTs, Weakly similar to 16.7Kd protein [1.00	115.00
	414577	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to	0.49	1.94
	414700	H63202	Hs.38163	ESTs	0.03	3.75
	415078	AA311223	Hs.283091	found in inflammatory zone 3	0.86	1.95
75	415120	N64464	Hs.34950	ESTs	1.00	120.00
	415323	BE269352	Hs.949	neutrophil cytosolic factor 2 (65kD, chr	0.60	2.48
	415335	AA847758	Hs.111030	ESTs	1.00	95.00
	415582	W92445	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT	1.00	136.00
	416030	H15261	Hs.21948	ESTs	0.02	8.07
80	416427	BE244050	Hs.79307	Rac/Cdc42 guanine exchange factor (GEF)	1.00	73.00
	416464	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	0.70	3.36
	416585	X54162	Hs.79386	leiomodulin 1 (smooth muscle)	0.06	6.56
	416847	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.70	3.65
	417148	AA359896	Hs.293885	hypothetical protein FLJ14902	1.00	114.00
85	417370	T28651	Hs.82030	tryptophanyl-tRNA synthetase	0.85	1.30
	417673	T87281	Hs.16355	ESTs	0.16	15.54

	418067	AI127958	Hs.83393	cystatin E/M	0.81	1.74
	418296	C01566	Hs.86671	ESTs	1.00	99.00
	418643	J03798	Hs.86948	small nuclear ribonucleoprotein D1 polyp	1.00	60.00
5	418832	X04011	Hs.88974	cytochrome b-245, beta polypeptide (chro	2.40	14.74
	418945	BE246762	Hs.89499	arachidonate 5-lipoxygenase	0.67	3.16
	419261	X07876	Hs.89791	wingless-type MMTV integration site fami	1.00	73.00
	419564	U08989	Hs.91139	solute carrier family 1 (neuronal/epitha	1.00	192.00
	419574	AK001989	Hs.91165	hypothetical protein	1.00	94.00
10	419968	X04430	Hs.93913	interleukin 6 (Interferon, beta 2)	61.16	500.00
	420266	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula	0.52	1.70
	420285	AA258124	Hs.293878	ESTs, Moderately similar to ZN91_HUMAN Z	1.00	172.00
	420577	AA278436	Hs.186649	ESTs	1.00	97.00
	421262	AA286746	Hs.9343	Homo sapiens cDNA FLJ14265 fis, clone PL	1.00	64.00
15	421445	AA913059	Hs.104433	Homo sapiens, clone IMAGE:4054868, mRNA	0.88	1.51
	421470	R27496	Hs.1378	annexin A3	0.05	11.26
	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	1.00	73.00
	421563	NM_006433	Hs.105806	granulysin	0.82	2.42
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	5.50	31.57
20	421855	F06504	Hs.27384	ESTs, Moderately similar to ALU4_HUMAN A	1.00	129.00
	421913	AI934365	Hs.109439	osteoglycin (osteoinductive factor, mime	1.00	101.00
	421952	AA300900	Hs.98849	ESTs, Moderately similar to AF161511 1 H	0.60	63.60
	422232	D43945	Hs.113274	transcription factor EC	1.00	148.00
	422386	AF105374	Hs.115830	heparan sulfate (glucosamine) 3-O-sulfot	1.40	3.98
25	423168	R34385	Hs.124940	GTP-binding protein	0.34	3.59
	423196	AK001866	Hs.125139	hypothetical protein FLJ11004	0.55	2.00
	423387	AJ012074		vasoactive intestinal peptide receptor 1	0.09	2.13
	423424	AF150241	Hs.128433	prostaglandin D2 synthase, hematopoietic	1.00	141.00
	423456	AL110151	Hs.128797	DKFZP586D0824 protein	1.00	66.00
30	423696	Z92546		Sushi domain (SCR repeat) containing	0.73	1.27
	424027	AW337575	Hs.201591	ESTs	0.54	2.58
	424212	NM_005814	Hs.143131	glycoprotein A33 (transmembrane)	0.77	2.47
	425087	R62424	Hs.126059	ESTs	1.00	74.00
	425175	AF020202	Hs.155001	UNC13 (C. elegans)-like	0.85	1.96
35	425771	BE561776	Hs.159494	Bruton agammaglobulinemia tyrosine kinas	1.18	2.56
	426486	BE178285	Hs.170056	Homo sapiens mRNA; cDNA DKFZp586B0220 (f	1.00	76.00
	427507	AF240467	Hs.179152	tol-like receptor 7	1.00	63.00
	427618	NM_000760	Hs.2175	colony stimulating factor 3 receptor (gr	0.60	2.19
	427732	NM_002980	Hs.2199	secretin receptor	0.97	1.42
40	427952	AA765368	Hs.293941	ESTs, Moderately similar to A53959 throm	1.00	105.00
	428709	BE268717	Hs.104916	hypothetical protein FLJ21940	1.00	80.00
	428769	AW207175	Hs.105771	ESTs	0.09	2.55
	428780	AI478578	Hs.50635	ESTs	1.00	98.00
	428833	AI928355	Hs.185805	ESTs	1.00	113.00
45	429657	D13626	Hs.2465	KIAA0001 gene product; putative G-protei	1.00	52.00
	430212	AA469153		gb:nc67f04.s1 NCI_CGAP_Pr1 Homo sapiens	1.00	132.00
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	0.11	15.60
	430376	AW292053	Hs.12532	chromosome 1 open reading frame 21	1.00	103.00
	430414	AW365665	Hs.120388	ESTs	0.50	6.96
50	430656	AA482900	Hs.162080	ESTs	1.00	70.00
	430843	AI734149	Hs.119514	ESTs	1.00	90.00
	430998	AF128847	Hs.204038	indolethylamine N-methyltransferase	0.29	1.84
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	1.00	79.00
	431921	N46466	Hs.58879	ESTs	0.91	1.67
55	432176	AW090386	Hs.112278	arrestin, beta 1	0.66	2.63
	432203	AA305748	Hs.49	macrophage scavenger receptor 1	1.00	76.00
	432231	AA339977	Hs.274127	CLST 11240 protein	0.46	1.46
	432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	0.79	2.25
	432522	D11466	Hs.51	phosphatidylinositol glycan, class A (pa	1.93	4.83
60	432596	AJ224741	Hs.278461	matrilin 3	0.04	5.79
	432850	X87723	Hs.3110	angiotensin receptor 2	1.00	167.00
	433138	AB029496	Hs.59729	semaphorin sem2	0.04	9.16
	433563	AI732637	Hs.277901	ESTs	1.00	91.00
	433588	AI056872	Hs.133386	ESTs	120.16	315.00
	434445	AI349306	Hs.11782	ESTs	0.60	1.84
65	435496	AW840171	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	128.00
	435974	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor	1.00	108.00
	436061	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	1.00	91.00
	437157	BE048860	Hs.120655	ESTs	1.00	87.00
70	437207	T27503	Hs.15929	hypothetical protein FLJ12910	1.00	105.00
	437311	AA370041	Hs.9456	SWI/SNF related, matrix associated, acti	1.00	71.00
	437439	H29796	Hs.269622	ESTs	1.00	115.00
	438199	AW016531	Hs.122147	ESTs	1.00	80.00
	439551	W72062	Hs.11112	ESTs	0.30	3.10
75	440515	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene fami	1.00	77.00
	440887	AI799488	Hs.135905	ESTs	1.00	85.00
	441025	AA913880	Hs.176379	ESTs	1.00	82.00
	441384	AA447849	Hs.288660	Homo sapiens cDNA: FLJ22182 fis, clone H	0.79	1.89
	441735	AI738675	Hs.127346	ESTs	1.00	75.00
80	442200	AW590572	Hs.235768	ESTs	0.78	5.83
	442832	AW206560	Hs.263569	ESTs	0.03	10.88
	442957	AI949952	Hs.49397	ESTs	1.00	70.00
	443282	T47764	Hs.132917	ESTs	1.00	197.00
	443547	AW271273	Hs.23767	hypothetical protein FLJ12666	1.00	253.00
85	443951	F13272	Hs.111334	ferritin, light polypeptide	0.55	2.09
	444330	AI597655	Hs.49265	ESTs	1.00	90.00

	444515	AW204908	Hs.169979	ESTs	1.00	84.00
	445769	AI741471	Hs.23666	ESTs	0.02	4.38
	445908	R13580	Hs.13436	Homo sapiens clone 24425 mRNA sequence	1.00	97.00
5	446291	BE397753	Hs.14623	interferon, gamma-inducible protein 30	0.93	1.69
	446917	AI347863	Hs.156672	ESTs	1.00	106.00
	447261	NM_006691	Hs.17917	extracellular link domain-containing 1	0.40	47.20
	447432	AW958473	Hs.301957	nudix (nucleoside diphosphate linked mol	1.00	100.00
	447482	AB033059	Hs.18705	KIAA1233 protein	0.05	8.21
10	447997	H00556	Hs.29792	ESTs, Weakly similar to I38022 hypotheti	0.02	5.42
	448299	AA497044	Hs.20887	hypothetical protein FLJ10392	1.00	79.00
	448782	AL050295	Hs.22039	KIAA0758 protein	0.42	1.56
	450575	NM_005859	Hs.29117	purine-rich element binding protein A	0.17	11.33
	450584	AA040403	Hs.60371	ESTs	1.00	94.00
15	450693	AW450461	Hs.203965	ESTs	1.00	91.00
	450715	AI266484	Hs.31570	ESTs, Weakly similar to KIAA1324 protein	1.00	152.00
	451103	R52804	Hs.25956	DKFZP564D206 protein	1.00	86.00
	451220	AF124251	Hs.26054	novel SH2-containing protein 3	0.60	1.30
	451668	Z43948	Hs.326444	cartilage acidic protein 1	0.54	1.91
20	452197	AW023595	Hs.232048	ESTs	1.00	67.00
	452331	AA598509	Hs.29117	purine-rich element binding protein A	4.53	11.07
	452353	C18825	Hs.29191	epithelial membrane protein 2	0.72	2.24
	453049	BE537217	Hs.30343	ESTs	1.00	68.00
	453107	NM_016113	Hs.279746	vanilloid receptor-like protein 1	0.83	1.70
25	453355	AW295374	Hs.31412	Homo sapiens cDNA FLJ11422 fis, clone HE	1.00	132.00
	453390	AA862496	Hs.28482	ESTs	1.00	72.00
	453531	AA417940		ESTs, Weakly similar to JC5795 CDEP prot	1.00	68.00
	454741	BE154396		gb:CM2-HT0342-091299-050-b05 HT0342 Homo	0.57	2.89
	456579	AA287827	Hs.284205	up-regulated by BCG-CWS	1.00	82.00
30	456672	AK020216	Hs.114727	Homo sapiens, clone MGC:16327, mRNA, com	0.79	1.96
	457400	AF032906	Hs.252549	cathepsin Z	1.03	3.25
	457718	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S	1.00	113.00
	459696	F03027		gb:HSC1KA072 normalized infant brain cDN	1.00	544.00

TABLE 10B

Pkey: Unique Eos probaset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

40	Pkey	CAT Number	Accession
	408074	103684_1	R20723 AA263003 AA333976 AA334725 AA334151 AW965490 AA310513 AI810530 D31302 AW134897 AA830127 AA046953 AI668930 C06094 AW104534
	411667	1253334_1	BE160198 AW935898 T11520 AW935930 AW856073 AW861034
45	413533	1375344_1	BE146973 BE146972 BE147042 BE147018 BE146783 BE147020 BE146781 BE147019 BE146766 BE147021 BE146952 BE146767 BE147044 BE146797 BE146776 BE146985 BE146793 BE146768 BE146771 BE146954 BE146760 BE147048 BE147025 BE147030
	423387	22779_1	AJ012074 U11087 L13288 X75299 L20295 AW630780 H14880 T28037 AI872991 R72136 AW449839 T81622 T79697 T29519 R94105 T83923 R73300 AI797007 R73390 AA961010 H74168 AI689932 BE045543 AI808418 AI608912 AI806573 AW884084 AW872978 AW872985 AA565655 AI022915 R50647 R73210 H45098 R46451 AW166269 T71132 AI264547 R52146 AI304920 R73391 AW884059 AW884085 H73241 T60038 T79612 R73145 R50549 AI094557 AI668793 R72302 AI564366 W01956 AA418962 W32571 R72840 H45409 R72085 R46356 R46758
50	423696	23112_1	AA508805 AA418798 T83751 R94072 T16182 AA928785 AA903896 Z92546 AA330586 AI570568 AW341487 AI827050 AW298668 AI792189 AI015693 AI733599 AI572251 AI672488 AW193262 AI244716 AI864375 AI206100 AA912444 AI269365 AI640254 AW772466 AI867336 AA627604 H16914 AA358477 AA338009
	430212	314437_1	AA469153 AI718503 AA469225
55	436532	421802_1	AA721522 AW975443 T93070
	453531	97026_1	AA417940 AA036735 T07025
	454741	1232559_1	BE154395 AW817959 BE154393

TABLE 10C

Pkey: Unique number corresponding to an Eos probaset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham L. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham L. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

	Pkey	Ref	Strand	Nt_position
70	400754	7331445	Plus	144559-144684
	401045	8117619	Plus	90044-90184,91111-91345
	401083	3242744	Plus	33192-33360
	402474	7547175	Minus	53526-53628,55755-55920,57530-57757
	402808	6456148	Minus	114964-115136,115461-115585,115931-116047,117666-117771,118004-118102
	403021	7547270	Plus	120799-120966
75	403421	9665041	Minus	126609-126773,139986-140205
	403438	9719679	Plus	90792-90938
	403687	7387384	Plus	9009-9534
	403764	7717105	Minus	118692-118853
	404277	1834458	Minus	91665-91946
80	404288	2769644	Plus	3512-3691
	404394	3135305	Minus	37121-37205,37491-37762,41053-41140,41322-41593,41773-41919
	404518	8151988	Plus	84494-84603
	404916	7341826	Plus	91057-91188
	405106	8079395	Minus	80877-81418
85	405257	7329310	Plus	73121-73273
	405381	6006920	Minus	7636-8054

TABLE 11A: Genes Distinguishing Adenocarcinoma from Other Lung Diseases and Normal Lung

Table 11A shows about 84 genes upregulated in lung adenocarcinomas relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 11B show the accession numbers for those Pkey's lacking UnigeneID's for table 11A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 11C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 11A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset Identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
403329			Target Exon	1.00	61.00
406399			NM_003122*:Homo sapiens serine protease	1.00	39.00
406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	226.37	350.00
407869	A827976	Hs.24391	hypothetical protein FLJ13612	0.77	1.18
407881	AW072003	Hs.40968	heparan sulfate (glucosamine) 3-O-sulfot	1.00	10.00
408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
409187	AF154830	Hs.50966	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
410076	T05387	Hs.7991	ESTs	1.12	1.50
410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
410399	BE068889		synuclein, gamma (breast cancer-specific	0.92	1.06
411908	L27943	Hs.72924	cytidine deaminase	1.00	1.00
412612	NM_000047	Hs.74131	arylsulfatase E (chondrodysplasia puncta	1.02	1.03
414075	U11862	Hs.75741	amidolite binding protein 1 (amine oxida	0.84	1.07
416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
417542	J04129	Hs.82269	progesterone-associated endometrial prole	1.28	1.35
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
419502	AU076704		fibrinogen, A alpha polypeptide	13.05	115.00
419631	AW186117	Hs.303154	popeye protein 3	1.00	13.00
420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
421190	U95031	Hs.102482	mucin 5, subtype B, tracheobronchial	1.17	1.55
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase 1, l	1.00	3.00
421582	A910275		trefoil factor 1 (breast cancer, estroge	1.23	1.00
422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
422095	A868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
423472	AF041260	Hs.129057	breast carcinoma amplified sequence 1	48.13	72.00
423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
424502	AF242388	Hs.149585	lengsin	1.00	1.00
424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	1.00	59.00
424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
424980	BE245380	Hs.153952	5' nucleotidase (CD73)	1.00	1.00
425523	AB007948	Hs.158244	KIAA0479 protein	1.00	35.00
426230	AA367019	Hs.241395	protease, serine, 1 (trypsin 1)	1.00	83.00
427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
428585	AB007863	Hs.185140	KIAA0403 protein	1.00	6.00
428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	1.59	1.69
430508	AJ015435	Hs.104637	ESTs	4.75	7.27
430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
431548	AJ834273	Hs.9711	novel protein	5.66	15.00
431566	AF176012	Hs.260720	J domain containing protein 1	49.76	37.00
431986	AA536130	Hs.149018	Novel human gene mapping to chromosome 20	1.19	1.47
432375	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
432677	NM_004482	Hs.278511	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
433819	AW511097	Hs.112765	ESTs	3.71	8.00
434001	AW950905	Hs.3697	serine (or cysteine) proteinase inhibito	29.31	72.00
434424	AJ811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
434792	AA649253	Hs.132458	ESTs	8.52	44.00
436217	TS3925	Hs.107	fibrinogen-like 1	57.97	31.00
436749	AA584890	Hs.5302	lectin, galactoside-binding, soluble, 4	1.10	1.41
436972	AA284679	Hs.25640	claudin 3	1.59	1.46
437866	AA156781		metallothionein 1E (functional)	3.62	101.00
437935	AW939591	Hs.5940	mucin 13, epithelial transmembrane	1.60	1.39
438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	1.00	1.00
439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00

	439759	AL359055	Hs.67709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00
	441377	BE218239	Hs.202656	ESTs	22.03	1.00
	443614	AV655386	Hs.7645	fibrinogen, B beta polypeptide	1.00	16.00
5	443813	AA876372	Hs.93951	Homo sapiens mRNA; cDNA DKFZp667D095 (fr	1.20	1.99
	443991	NM_002250	Hs.10082	potassium intermediate/small conductance	5.71	6.87
	444670	H58373	Hs.332938	hypothetical protein MGC5370	1.98	38.00
	444931	AV652056	Hs.75113	general transcription factor IIIA	1.00	54.00
	446102	AW168067	Hs.317694	ESTs	1.00	1.00
10	446163	AA026880	Hs.25252	Homo sapiens cDNA FLJ13603 fis, clone PL	1.00	36.00
	446469	BE094848	Hs.15113	homogenisate 1,2-dioxygenase (homogeni	1.00	11.00
	447388	AW630534	Hs.76277	Homo sapiens, clone MGC:9381, mRNA, comp	1.24	1.16
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
	448243	AW369771	Hs.52620	integrin, beta 8	15.84	1.00
15	448844	AI581519	Hs.177164	ESTs	1.00	31.00
	449444	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	1.00	83.00
	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	452889	F33868	Hs.284176	transferrin	1.54	1.44
	453392	U23762	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00
20	453464	AI884911	Hs.32989	receptor (calcitonin) activity modifying	1.55	2.45
	453735	AI066629	Hs.125073	ESTs	1.01	1.30

TABLE 11B

25 Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

	Pkey	CAT Number	Accession
30	410399	11995_1	BE068889 BE068882 AF044311 AF017256 NM_003087 AF037207 AF010126 AA633976 AA872836 BE298825 BE298889 AI016464 AI684600 AI936527 AA804675 AA394097 AI139933 AA946606 BE117133 AA722407 AA293803 AI688480 AA056035 AA055968 AW796957 AI637713 AA410737 H49348 AA488472 AA411094 AA235594 AA402824 AA443638 AW452137 AA421708 AW265211 AI493266 AA365132 AW966044 AU076704 T74854 T74860 T72098 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72826 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M54982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 R05795 H48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AA011465 AA345378 AV654847 AV654272 AV656001 AI064740 T82897 N33594 AA344542 AW805054 AI207457 T61743 AA026737 H94389 AA382695 AA918409 T68044 S82092 T39959 AI017721 AA312395 AA312919 T40156 H66239 AV652989 H38728 R98521 AV655200 R95790 W03250 W00913 AA344136 AV660126 R97823 AA343596 AW470774 AV651256 N54417 AA812862 AW182928 AI111192 H61463 H72060 AA344503 H38539 AI277511 AV661108 AI207625 T47810 AA235252 T7853 T47778 R95746 H70620 AA701463 AW827166 R98475 C20925 AV657287 T71959 T71313 T73920 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T7854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 AI110639 AA344603 AF063513 T64686 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74577 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA342807 AA341964 T53747 T72042 T62764 AI064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 AI207557 T68219 W86031 T69081 T64232 R93198 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 AI248502 R29454 T64764 T57001 T73052 T71429 T51176 T58866 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 AI910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 AI571948 AA507595 AA614579 AA587613 R83818 AA568312 AA614409 AA307578 AI925552 AW950155 AI910083 M12075 BE074052 AW004668 AA578674 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA858276 R56389 AI001051 AW050700 AW750216 AA614539 BE074045 AI307407 AW802303 BE073575 AI202532 AA524242 AI970839 AI909751 BE076078 AI509749 R55292 AA156781 AW293839 U52054 AA024963 AT778446 BE073977 AW444904 AW602574 BE164040 BE164012 BE163972 BE163974 BE163992 AA837481 AW468444 BE185091 AW468002 AA687333 AA811830 AA581806 AI866686 AI572124 AA043777 AA040926 D20160 AI536733 AA812489 AW874142 AI471883 W84421 AA156850 W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW450652 AW449519 AA993634 AI806539 AA351618 AW449522 AI827628 AA904788 AA380381 AA886045 AA774409 BE003229 Z41756

TABLE 11C

70 Pkey: Unique number corresponding to an Eos probeset
Ref: Sequence source. The 7 digit numbers in this column are Genbank identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand: Indicates DNA strand from which exons were predicted.
Nt_position: Indicates nucleotide positions of predicted exons.

	Pkey	Ref	Strand	Nt_position
75	403329	8516120	Plus	96450-96598
	406399	9256288	Minus	63448-63554

TABLE 12A: Genes Distinguishing Squamous Cell Carcinoma from Other Lung Diseases and Normal Lung

Table 12A shows about 72 genes upregulated in squamous cell carcinomas of the lung relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 12B show the accession numbers for those Pkey's lacking UnigeneID's for table 12A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwin, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 12C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 12A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400666			NM_002425:Homo sapiens matrix metallopro	3.26	3.22
401780			NM_005557*:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275*:Homo sapiens keratin 15 (KRT1	4.13	2.70
401994			Target Exon	61.84	47.00
402075			ENSP00000251056*:Plasma membrane calcium	1.00	1.00
404996			Target Exon	1.00	1.00
407839	AA045144	Hs.161566	ESTs	173.91	108.00
408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	151.17	8.00
408522	A1541214	Hs.46320	Small proline-rich protein SPRK (human,	1.98	1.24
410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
415817	U98967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
417366	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)	8.97	3.27
418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
420783	A1659838	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
421773	W69233	Hs.112457	ESTs	1.12	1.14
421948	L42583	Hs.334309	keratin 6A	51.83	20.25
421978	AJ243662	Hs.110196	NICE-1 protein	1.01	0.91
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	233.42	68.00
424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor	1.00	1.00
424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00
427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	11.80	1.00
430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
430890	X54232	Hs.2699	glypican 1	1.58	1.40
431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
431846	BE019924	Hs.271580	uropodxin 1B	4.49	2.51
433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
434360	AW015415	Hs.127780	ESTs	40.98	27.00
434880	U02388	Hs.101	cytochrome P450, subfamily IVF, polypept	1.00	1.00
435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
436511	AA721252	Hs.291502	ESTs	16.76	14.00
438403	AA806607	Hs.292206	ESTs	1.00	1.00
439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
439605	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
439706	AW872527	Hs.59761	ESTs, Weakly similar to DAPI_HUMAN DEATH	86.55	11.00
440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
441525	AW241867	Hs.127728	ESTs	1.53	1.42
443162	T49951	Hs.9029	DKFZP434G032 protein	31.11	39.00
444378	R41339	Hs.12569	ESTs	1.00	1.00

WO 02/086443

PCT/US02/12476

	446292	AF081497	Hs.279682	Rh type C glycoprotein	1.55	1.26
	447078	AW885727	Hs.9914	ESTs	47.24	24.00
	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
5	449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00
	450832	AW970602	Hs.105421	ESTs	25.17	36.00
	452240	AI591147	Hs.61232	ESTs	13.42	1.00
	453317	NM_002277	Hs.41696	keratin, hair, acidic, 1	1.19	1.27
	453830	AA534296	Hs.20953	ESTs	24.92	25.00
10	454099	W27553	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11
	455601	AI368680	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00

TABLE 12B

15	Pkey:	Unique Eos probeset Identifier number				
	CAT number:	Gene cluster number				
	Accession:	Genbank accession numbers				
20	Pkey	CAT Number	Accession			
	439285	47065_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955694 AI346341 AI867454 N54784 AI655270 AI421279 AW014882 AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077			

TABLE 12C

25	Pkey:	Unique number corresponding to an Eos probeset				
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.				
30	Strand:	Indicates DNA strand from which exons were predicted.				
	NL_position:	Indicates nucleotide positions of predicted exons.				
	Pkey	Ref	Strand	NL_position		
	400666	8118498	Plus	17982-18115,20297-20456		
	401780	7249190	Minus	28397-28617,28920-29045,29135-29296,29411-29567,29705-29787,30224-30573		
35	401781	7249190	Minus	83215-83435,83531-83656,83740-83901,84237-84393,84955-85037,85290-86814		
	401785	7249190	Minus	165776-165998,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942		
	401994	4153858	Minus	42904-43124,43211-43338,44607-44763,45199-45281,46337-46732		
	402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076		
40	404996	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40674,42351-42450		

TABLE 13A: Genes Distinguishing Non-Malignant Lung Disease from Lung Tumors and Normal lung

Table 13A shows about 23 genes upregulated in non-malignant lung disease relative to lung tumors and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 13B show the accession numbers for those Pkey's lacking UnigenelD's for table 13A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 13C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 13A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigenel number
 Unigenel Title: Unigenel gene title
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigenelD	Unigenel Title	R1	R2
408562	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
409031	AA376836	Hs.76728	ESTs	1.00	128.00
412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00
417511	AL049176	Hs.82223	chordin-like	1.00	179.00
418819	AA228776	Hs.191721	ESTs	1.00	140.00
422060	R20893	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.00	156.00
424585	AA464840	Hs.131987	ESTs	1.00	167.00
426753	T89832	Hs.170278	ESTs	1.00	141.00
429496	AA453800	Hs.192793	ESTs	1.00	138.00
430719	AA488988	Hs.293796	ESTs	1.00	133.00
431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
431728	NM_007351	Hs.268107	multimerin	1.00	157.00
436532	AA721522		gbmv54h12.r1 NCL_CGAP_Ew1 Homo sapiens	1.00	218.00
437960	AI669586	Hs.222194	ESTs	1.00	147.00
438202	AW169287	Hs.22588	ESTs	1.00	141.00
441499	AW298235	Hs.101689	ESTs	1.00	167.00
444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
448253	H25899	Hs.201591	ESTs	1.00	141.00
453636	R57837	Hs.169872	ESTs	1.00	116.00
458332	AI000341	Hs.220491	ESTs	1.00	192.00
459587	AA031956		gbzkl5e04.s1 Soares_pregnant_uterus_NbH	1.00	154.00

TABLE 13B

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
431089	327825_1	BE041395 AA491826 AA621946 AA715980 AA666102
436532	421802_1	AA721522 AW975443 T93070

TABLE 13C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 NL_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076

TABLE 14A: Preferred Utility and Subcellular Localization for Potential Lung Disease Targets

Table 14A shows the subcellular localization and preferred utility for the genes appearing in Tables 9A and 10A. mAb symbolizes monoclonal antibody, diag symbolizes diagnostic, s.m. symbolizes small molecule, and CTL symbolizes cytotoxic lymphocytic ligand. These genes were selected from 59680 probesets on the Eos/Affymatrix Hu03 Genechip array.

Table 14B show the accession numbers for those Pkey's lacking UnigenelD's for table 14A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 14C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 14A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number

ExAccn: Exemplar Accession number, Genbank accession number

UnigenelD: Unigenel number

Unigenel Title: Unigenel gene title

Pref.Utility: Preferred Utility

Pred.Loc: Predicted subcellular localization

Pkey	ExAccn	UnigenelD	Unigenel Title	Pref Utility	Pred. Loc
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	mAb & diag & s.m.	extracellular
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	mAb	plasma membrane
402075			ENSP00000251056:Plasma membrane calcium	mAb & diag	secreted
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagonist	diag	secreted
408243	Y00787	Hs.624	interleukin 8	diag	secreted
408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	mAb & s.m.	plasma membrane
408908	BE296227	Hs.250822	serine/threonine kinase 15	s.m.	cytoplasm
409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	CTL & diag	secreted
409103	AF251237	Hs.112208	XAGE-1 protein	CTL	nuclear
409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	diag	secreted
409632	V74001	Hs.55279	serine (or cysteine) proteinase inhibitor	diag	secreted
409757	NM_001898	Hs.123114	cystatin SN	diag	extracellular
409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	CTL	nuclear
409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	diag	extracellular
410001	AB041036	Hs.57771	kallikrein 11	diag	extracellular
410407	X66839	Hs.63287	carbonic anhydrase IX	mAb & s.m.	plasma membrane
410418	D31382	Hs.63325	transmembrane protease, serine 4	mAb & diag & s.m.	plasma membrane
412140	AA219591	Hs.73625	RAB6 interacting, kinesin-like (rabkines	s.m.	
412719	AWD16610	Hs.816	ESTs	s.m.	nuclear
414774	X02419	Hs.77274	plasminogen activator, urokinase	diag	extracellular
414883	AA926960		CDC28 protein kinase 1	s.m.	
415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	CTL & diag	extracellular
415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	mAb & diag & s.m.	secreted
415817	U89967	Hs.78867	protein tyrosine phosphatase, receptor-t	mAb & s.m.	plasma membrane
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	diag	extracellular
417034	NM_006183	Hs.80962	neurotensin	diag	extracellular
417079	U65590	Hs.81134	interleukin 1 receptor antagonist	diag	extracellular
417308	H60720	Hs.81892	KIAA0101 gene product	s.m.	mitochondrial
417389	BE260964	Hs.82045	midkine (neurile growth-promoting factor	mAb & diag	secreted
417433	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein	mAb	plasma membrane
417933	X02308	Hs.82962	thymidylate synthetase	s.m.	endoplasmic reticulum
418478	U98945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	s.m.	cytoplasm
418506	AA084248	Hs.85339	G protein-coupled receptor 39	mAb & s.m.	plasma membrane
418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)	CTL	cytoplasmic
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	diag	secreted
419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor t	mAb & s.m.	plasma membrane
419183	U60689	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	CTL & s.m.	mitochondrial
419216	AU078718	Hs.164021	small inducible cytokine subfamily B (Cy	diag	secreted
419235	AW470411	Hs.288433	neurotrophin	mAb & diag	plasma membrane
419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	mAb & s.m.	plasma membrane
419556	U29615	Hs.91093	chitinase 1 (chitinotrioidase)	mAb & diag	extracellular
420610	AJ683183	Hs.99348	distal-less homeo box 5	CTL	nuclear
421110	AJ250717	Hs.1355	cathepsin E	sm & diag	extracellular
421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	diag	secreted
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	mAb & s.m.	plasma membrane
421552	AF026592	Hs.105700	secreted frizzled-related protein 4	diag	secreted
421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	mAb & s.m.	plasma membrane
421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	mAb & s.m.	plasma membrane
422109	S73265	Hs.1473	gastrin-releasing peptide	diag	secreted
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	diag	secreted
422282	AF019225	Hs.114309	apolipoprotein L	diag	secreted
422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.carevis	s.m.	nuclear
422424	AJ186431	Hs.296638	prostate differentiation factor	diag	extracellular
422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	s.m.	cytoplasm
422809	AK001379	Hs.121028	hypothetical protein FLJ10549	s.m.	nuclear
422867	L32137	Hs.1584	collagen oligomeric matrix protein (ose	diag	extracellular
422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	CTL & s.m.	
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	diag	
423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	mAb & diag & s.m.	secreted
423961	D13666	Hs.136348	periostin (OSF-2os)	mAb & diag	extracellular
424046	AF027856	Hs.138202	serine (or cysteine) proteinase inhibitor	diag	secreted
424381	AA285249	Hs.146329	protein kinase Ctk2	s.m.	nuclear

	424502	AF242388	Hs.149585	lengsin	s.m.	cytoplasmic
	424503	NM_002205	Hs.149609	integrin, alpha 5 (fibronectin receptor,	mAb & s.m.	plasma membrane
	424587	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	diag	extracellular
5	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	mAb & diag & s.m.	secreted
	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	s.m.	cytoplasmic
	425650	NM_001944	Hs.1925	desmoglein 3 (periphagus vulgaris antigen	mAb	plasma membrane
	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	s.m.	
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	mAb & diag	plasma membrane
10	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	mAb & s.m.	plasma membrane
	426215	AW963419	Hs.155223	stanniocalcin 2	mAb & diag	secreted
	426427	M86699	Hs.169840	TTK protein kinase	CTL & s.m.	nuclear
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	mAb & diag	secreted
	427335	AA448542	Hs.251677	G antigen 7B	CTL	cytoplasmic
	427747	AW411425	Hs.180655	serine/threonine kinase 12	s.m.	cytoplasmic
15	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	diag	
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	mAb & diag & s.m.	extracellular
	428450	NM_014791	Hs.184339	KIAA0175 gene product	s.m.	nuclear
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	s.m.	nuclear
20	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	mAb & s.m.	plasma membrane
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	CTL & s.m.	nuclear
	428698	AA852773	Hs.334838	KIAA1866 protein	mAb	
	428748	AW593205	Hs.98785	Ksp37 protein	diag	extracellular
	428758	AA433988	Hs.98502	CA125 antigen; mucin 16	diag	mitochondria*
	428969	AF102074	Hs.194689	artemin	diag	extracellular
25	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	mAb & s.m.	plasma membrane
	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	mAb & s.m.	plasma membrane
	429547	AW009166	Hs.99376	ESTs	diag	secreted
	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	mAb & diag	secreted
30	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	s.m.	
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	mAb & s.m.	plasma membrane
	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	diag	extracellular
	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	mAb & s.m.	plasma membrane
	431846	BE019924	Hs.271580	uroplakin 1B	mAb & diag	plasma membrane
35	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	mAb & diag	plasma membrane
	432201	AI538613	Hs.298241	Transmembrane protease, serine 3	mAb & diag & s.m.	plasma membrane
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	s.m.	nuclear
	435505	AF200492	Hs.211238	interleukin-1 homolog 1	diag	secreted
	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	s.m.	
40	437016	AU076916	Hs.5398	guanine monophosphate synthetase	s.m.	cytoplasm
	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an	CTL	ER
	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheri	CTL	nuclear
	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365012.1 [H.s.a	mAb & s.m.	plasma membrane
	439223	AW238299	Hs.250618	UL16 binding protein 2	mAb	plasma membrane
45	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	mAb & s.m.	
	439606	W79123	Hs.58551	G protein-coupled receptor 87	mAb & s.m.	plasma membrane
	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	mAb & s.m.	plasma membrane
	440006	AK000517	Hs.6844	NALP2 protein; PYRIN-Containing APAF1-I	s.m.	nuclear
	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Ra	s.m.	
50	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447	mAb & s.m.	plasma membrane
	443247	BE614387	Hs.333893	c-Myc target JPO1	CTL	extracellular*
	443426	AF098158	Hs.9329	chromosome 20 open reading frame 1	CTL	
	443859	NM_013409	Hs.9914	folistatin	diag	extracellular
	444006	BE395085	Hs.10086	type I transmembrane protein Fn14	mAb	plasma membrane
55	444371	BE540274	Hs.239	forkhead box M1	s.m.	nuclear
	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheri	diag	secreted
	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	mAb & diag	plasma membrane
	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6	mAb & diag	secreted
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	diag	secreted
60	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	diag	extracellular
	447033	AI357412	Hs.157601	ESTs	CTL & diag	secreted
	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	CTL	
	448243	AW369771	Hs.52620	integrin, beta 8	mAb & s.m.	plasma membrane
	448844	AI581519	Hs.177164	ESTs	mAb & s.m.	
65	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	mAb	plasma membrane
	449722	BE280074	Hs.23960	cyclin B1	s.m.	cytoplasm
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte	mAb & s.m.	plasma membrane
	450375	AA009647		a disintegrin and metalloproteinase doma	mAb & diag & s.m.	plasma membrane
	450701	H39960	Hs.288457	hypothetical protein XP_098151 (leucine-	mAb & diag	plasma membrane
70	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	diag	secreted
	451668	Z43948	Hs.326444	cartilage acidic protein 1	mAb & diag	plasma membrane
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	diag	
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	diag	extracellular
	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR	mAb	plasma membrane
	452838	U65011	Hs.30743	preferentially expressed antigen in mela	CTL	nuclear
75	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso	CTL & s.m.	nuclear
	457489	AI693815	Hs.127179	cryptic gene	diag	secreted

TABLE 148

80
 Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey CAT Number Accession

WO 02/086443

PCT/US02/12476

414883 15024_1

AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245
 AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW960304 T29812 AA476873 BE297387
 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 A1719075 BE270172 BE269819 AA889955 A1204630 W25243 A1935150
 AA872039 W72395 T99630 A1422691 H98460 N31428 BE255916 H03265 A1857576 AA776920 AA910644 AA459522 AA293140 AW514667
 R75953 AW662396 AA662522 A1865147 A1423153 AW262230 AA584410 AA583187 AW024595 AW069734 A1828996 AA282997 AA876046
 AW613002 AA527373 AW972459 A1831360 AA621337 AA100926 AA772418 AA594628 A1033892 W95096 A1034317 AA398727 A1085031
 N95210 A1459432 A1041437 AA932124 AA627684 AA935829 A1004827 A1423513 A1094597 H42079 R54703 A1630359 AA617681 AA978045
 AA643280 W44561 A1991988 A1537692 A1090262 AA740817 A1312104 A1911822 AA416871 A1185409 AA129784 AA701623 A1075239
 A1139549 AA633648 A1339996 A1336880 AA399239 A1078708 A1085351 A1362835 A1346618 A1146955 A1989360 A1348243 N92892 AA765850
 A1494230 A1278887 AA962596 A1492600 W80435 AA001979 R97424 A1129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785
 A1494211 AW059601 AW886710 R92790 N59755 A1361128 AW589407 H47725 H97534 H48076 H48450 T99631 AW300758 H03431 R76789
 AA954344 H77576 R96823 A1457100 N92845 N49682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923
 H03266 BE261918 AA769633 AA480310 AA507454 A910586 A1203723 AW104725 W25611 W25071 T88980 H03513 T77589 R99155
 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283672
 AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04394 W46291 AW663674 H04021 H01532
 AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 A1161014 AA099554 R69067

TABLE 14C

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 NL_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076

TABLE 15A: Information for all sequences in Table 16

Table 15A shows the Seq ID No, Pkey, ExAccn, UnigenelD, and Unigene Title for all of the sequences in Table 16.

5 Table 15B show the accession numbers for those Pkey's lacking UnigenelD's for table 15A. For each probaset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwin, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10 Table 15C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 15A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

15	Seq ID No:	Sequence ID number			
	Pkey:	Unique Eos probaset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigenelD:	Unigene number			
	Unigene Title:	Unigene gene title			
20	Seq ID No:	Pkey	ExAccn	UnigenelD	Unigene Title
	Seq ID No: 1 & 2	410407	X66839	Hs.63287	carbonic anhydrase IX
	Seq ID No: 3 & 4	412719	AW016610	Hs.816	ESTs
	Seq ID No: 5 & 6	417034	NM_006183	Hs.80962	neurotensin
25	Seq ID No: 7 & 8	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
	Seq ID No: 9 & 10	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
	Seq ID No: 11 & 12	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
	Seq ID No: 13 & 14	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
	Seq ID No: 15 & 16	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
30	Seq ID No: 17 & 18	439285	AL133916		hypothetical protein FLJ20093
	Seq ID No: 19 & 20	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kafirin
	Seq ID No: 21 & 22	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog
	Seq ID No: 23 & 24	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
	Seq ID No: 25 & 26	412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (raklines
35	Seq ID No: 27 & 28	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
	Seq ID No: 29 & 30	452838	U65011	Hs.30743	preferentially expressed antigen in mela
	Seq ID No: 31 & 32	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 33 & 34	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 35 & 36	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibito
40	Seq ID No: 37 & 38	429510	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
	Seq ID No: 39 & 40	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
	Seq ID No: 41 & 42	431846	BE019924	Hs.271580	uropod 1B
	Seq ID No: 43 & 44	418830	BE513731	Hs.88959	hypothetical protein MGC4816
	Seq ID No: 45 & 46	424098	AF077374	Hs.139322	small proline-rich protein 3
45	Seq ID No: 47 & 48	443648	AJ085377	Hs.143610	ESTs
	Seq ID No: 49	311034	BE567130	Hs.311369	ESTs, Highly similar to NKGD_HUMAN NKG2-
	Seq ID No: 50 & 51	408522	AJ541214	Hs.46320	Small proline-rich protein SPRK (human,
	Seq ID No: 52 & 53	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
	Seq ID No: 54 & 55	435505	AF200492	Hs.211238	interleukin-1 homolog 1
50	Seq ID No: 56 & 57	417366	BE185269	Hs.1076	small proline-rich protein 1B (comfalin)
	Seq ID No: 58 & 59	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
	Seq ID No: 60 & 61	441020	W79283	Hs.35962	ESTs
	Seq ID No: 62 & 63	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys
	Seq ID No: 64 & 65	429538	BE182592	Hs.11261	small proline-rich protein 2A
55	Seq ID No: 66 & 67	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte
	Seq ID No: 68 & 69	444371	BE540274	Hs.239	forkhead box M1
	Seq ID No: 70 & 71	444371	BE540274	Hs.239	forkhead box M1
	Seq ID No: 72 & 73	444371	BE540274	Hs.239	forkhead box M1
	Seq ID No: 74 & 75	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
60	Seq ID No: 76 & 77	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
	Seq ID No: 78 & 79	429259	AA420450	Hs.292911	Plakophilin
	Seq ID No: 80 & 81	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu
	Seq ID No: 82 & 83	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
	Seq ID No: 84 & 85	423662	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro
65	Seq ID No: 86 & 87	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
	Seq ID No: 88 & 89	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
	Seq ID No: 90 & 91	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor
	Seq ID No: 92 & 93	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr
	Seq ID No: 94 & 95	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated
70	Seq ID No: 96 & 97	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re
	Seq ID No: 98 & 99	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic
	Seq ID No: 100 & 101	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o
	Seq ID No: 102 & 103	431009	BE149762	Hs.48958	gap junction protein, beta 6 (connexin 3
	Seq ID No: 104 & 105	409103	AF251237	Hs.112208	XAGE-1 protein
75	Seq ID No: 106 & 107	417542	J04129	Hs.82269	progesterone-associated endometrial prote
	Seq ID No: 108 & 109	428471	X57348	Hs.184510	stratiferin
	Seq ID No: 110 & 111	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member
	Seq ID No: 112 & 113	414761	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2
	Seq ID No: 114 & 115	418203	X54942	Hs.83758	CDC28 protein kinase 2
80	Seq ID No: 116	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m
	Seq ID No: 117 & 118	437016	AU076916	Hs.5398	guanine monophosphate synthetase
	Seq ID No: 119 & 120	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
	Seq ID No: 121 & 122	446989	AK001898	Hs.16740	hypothetical protein FLJ11035
	Seq ID No: 123 & 124	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein
85	Seq ID No: 125 & 126	424687	J05070	Hs.161738	matrix metalloproteinase 9 (gelatinase B

	Seq ID No: 127 & 128	414430	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1
	Seq ID No: 129 & 130	418462	BE001596	Hs.85266	integrin, beta 4
	Seq ID No: 131 & 132	100668	L05424	Hs.169610	CD44 antigen (homing function and Indian
5	Seq ID No: 133 & 134	458933	AI638429	Hs.24763	RAN binding protein 1
	Seq ID No: 135 & 136	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 137 & 138	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 139 & 140	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 141 & 142	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
10	Seq ID No: 143 & 144	446269	AW263155	Hs.14559	hypothetical protein FLJ10540
	Seq ID No: 145 & 146	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur
	Seq ID No: 147 & 148	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con
	Seq ID No: 149 & 150	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma
	Seq ID No: 151 & 152	439606	W79123	Hs.58561	G protein-coupled receptor 87
15	Seq ID No: 153 & 154	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 155 & 156	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 157 & 158	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 159 & 160	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 161 & 162	404877			NM_005365: Homo sapiens melanoma antigen,
	Seq ID No: 163 & 164	413129	AF292100	Hs.104613	RP42 homolog
20	Seq ID No: 165 & 166	413281	AA861271	Hs.222024	transcription factor BMAL2
	Seq ID No: 167 & 168	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 169 & 170	416819	U77735	Hs.80205	pim-2 oncogene
	Seq ID No: 171 & 172	451320	AW118072		diacylglycerol kinase, zeta (104kD)
	Seq ID No: 173 & 174	418543	NM_005329	Hs.85962	hyaluronan synthase 3
25	Seq ID No: 175 & 176	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member
	Seq ID No: 177 & 178	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)
	Seq ID No: 179 & 180	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 181 & 182	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
30	Seq ID No: 183 & 184	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 185 & 186	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 187 & 188	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 189 & 190	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 191 & 192	448993	AI471630	Hs.8127	KIAA0144 gene product
35	Seq ID No: 193 & 194	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 195 & 196	430393	BE185030	Hs.241305	estrogen-responsive B box protein
	Seq ID No: 197 & 198	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol
	Seq ID No: 199 & 200	420462	AF050147	Hs.97932	chondromodulin I precursor
	Seq ID No: 201 & 202	102963	X02404	Hs.274534	calcitonin-related polypeptide, beta
40	Seq ID No: 203 & 204	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid
	Seq ID No: 205 & 206	101175	U82671	Hs.36980	melanoma antigen, family A, 2
	Seq ID No: 207 & 208	429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like
	Seq ID No: 209 & 210	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 211 & 212	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
45	Seq ID No: 213 & 214	131927	AJ003112	Hs.34780	doublecortin; lissencephaly, X-linked (d
	Seq ID No: 215 & 216	428182	BE386042	Hs.293317	ESTs, Weady similar to GGC1_HUMAN G ANT
	Seq ID No: 217 & 218	427335	AA448542	Hs.251677	G antigen 7B
	Seq ID No: 219 & 220	409420	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kalini
	Seq ID No: 221 & 222	114346	AL137256	Hs.130489	ATPase, aminophospholipid transporter-II
50	Seq ID No: 223 & 224	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9
	Seq ID No: 225 & 226	404440			NM_021048: Homo sapiens melanoma antigen,
	Seq ID No: 227 & 228	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor
	Seq ID No: 229 & 230	103312	Y12642	Hs.3185	lysosomal
	Seq ID No: 231 & 232	320843	BE069288	Hs.34744	Homo sapiens mRNA: cDNA DKFZp547C136 (fr
55	Seq ID No: 233	429065	AJ753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT
	Seq ID No: 234 & 235	446102	AW168067	Hs.317694	ESTs
	Seq ID No: 236 & 237	330495	U47924	Hs.71642	guanine nucleotide binding protein (G pr
	Seq ID No: 238	413573	AI733859	Hs.149089	ESTs
	Seq ID No: 239 & 240	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
60	Seq ID No: 241 & 242	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
	Seq ID No: 243 & 244	332180	AF134160	Hs.7327	claudin 1
	Seq ID No: 245	437915	AI637993	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca
	Seq ID No: 246 & 247	441553	AA281219	Hs.121296	ESTs
	Seq ID No: 248 & 249	331692	AI683487	Hs.152213	wingless-type MMTV integration site fami
65	Seq ID No: 250 & 251	429413	NM_014058	Hs.201877	DESC1 protein
	Seq ID No: 252 & 253	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis
	Seq ID No: 254 & 255	448357	N20169	Hs.108923	RAB38, member RAS oncogene family
	Seq ID No: 256 & 257	446292	AF081497	Hs.279682	Rh type C glycoprotein
	Seq ID No: 258 & 259	416209	AA235776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h
70	Seq ID No: 260 & 261	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1
	Seq ID No: 262 & 263	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor
	Seq ID No: 264 & 265	439223	AW238299	Hs.250618	UL16 binding protein 2
	Seq ID No: 266 & 267	429228	AI553633	Hs.326447	ESTs
	Seq ID No: 268 & 269	409757	NM_001898	Hs.123114	cystatin SN
75	Seq ID No: 270 & 271	411089	AA456454	Hs.214291	cell division cycle 2-like 1 (PITSLRE pr
	Seq ID No: 272 & 273	436511	AA721252	Hs.291502	ESTs
	Seq ID No: 274 & 275	428969	AF120274	Hs.194689	artemin
	Seq ID No: 276 & 277	428969	AF120274	Hs.194689	artemin
	Seq ID No: 278 & 279	428969	AF120274	Hs.194689	artemin
80	Seq ID No: 280 & 281	428969	AF120274	Hs.194689	artemin
	Seq ID No: 282	407137	T97307		gb:ye53h05.s1 Soares fetal liver spleen
	Seq ID No: 283 & 284	412723	AA648459	Hs.335951	hypothetical protein AF301222
	Seq ID No: 285 & 286	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 287 & 288	405770			NM_002362: Homo sapiens melanoma antigen,
85	Seq ID No: 289 & 290	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13
	Seq ID No: 291 & 292	414774	X02419	Hs.77274	plasminogen activator, urokinase

	Seq ID No: 293 & 294	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub
	Seq ID No: 295 & 296	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 297 & 298	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
5	Seq ID No: 299 & 300	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 301 & 302	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 303 & 304	437789	AJ581344	Hs.127812	ESTs, Weakly similar to T17330 hypothe
	Seq ID No: 305 & 306	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso
	Seq ID No: 307 & 308	403478			NM_022342:Homo sapiens kinesin protein 9
	Seq ID No: 309	441525	AW241867	Hs.127728	ESTs
10	Seq ID No: 310 & 311	434105	AW952124	Hs.13094	presenilins associated rhomboid-like pro
	Seq ID No: 312 & 313	428810	AF058236	Hs.193788	nitric oxide synthase 2A (inducible, hep
	Seq ID No: 314 & 315	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B
	Seq ID No: 316 & 317	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f
	Seq ID No: 318 & 319	409228	R16811	Hs.22010	ESTs, Weakly similar to 2109260A B cell
15	Seq ID No: 320 & 321	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg
	Seq ID No: 322 & 323	413582	AW295647	Hs.71331	hypothetical protein MGC5350
	Seq ID No: 324 & 325	438403	AA806607	Hs.292206	ESTs
	Seq ID No: 326 & 327	403329			unnamed protein product [Homo sapiens]
	Seq ID No: 328 & 329	409893	AW247090	Hs.57101	mitochondrion maintenance deficient (S.
20	Seq ID No: 330 & 331	411973	BE245360	Hs.279477	v-ets erythroblastosis virus E26 oncogen
	Seq ID No: 332 & 333	413195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	Seq ID No: 334 & 335	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	Seq ID No: 336 & 337	101345	NM_005795	Hs.152175	calcitonin receptor-like
	Seq ID No: 338 & 339	103280	U84722	Hs.76205	cadherin 5, type 2, VE-cadherin (vascula
25	Seq ID No: 340 & 341	102012	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
	Seq ID No: 342 & 343	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	Seq ID No: 344 & 345	134299	AW580939	Hs.97199	complement component C1q receptor
	Seq ID No: 346 & 347	412719	AW016610	Hs.816	ESTs
30	Seq ID No: 348 & 349	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
	Seq ID No: 350 & 351	128924	BE279383	Hs.26557	plakophilin 3
	Seq ID No: 352 & 353	100486	T19006	Hs.10842	RAN, member RAS oncogene family
	Seq ID No: 354 & 355	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 356 & 357	409459	D86407	Hs.54481	low density lipoprotein receptor-related
	Seq ID No: 358 & 359	330493	M27826		endogenous retroviral protease
35	Seq ID No: 360 & 361	417866	AW067903	Hs.82772	collagen, type XI, alpha 1
	Seq ID No: 362 & 363	418113	AJ272141	Hs.83484	SRY (sex determining region Y)-box 4
	Seq ID No: 364 & 365	437016	AU076916	Hs.5398	guanine monophosphate synthetase
	Seq ID No: 366 & 367	429612	AF062649	Hs.252587	pituitary tumor-transforming 1
40	Seq ID No: 368 & 369	440704	M69241	Hs.162	insulin-like growth factor binding prote
	Seq ID No: 370 & 371	431221	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye
	Seq ID No: 372 & 373	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 374 & 375	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 376 & 377	132354	BE185289	Hs.1076	small proline-rich protein 1B (comiflin)
45	Seq ID No: 378 & 379	424441	X14850	Hs.147097	H2A histone family, member X
	Seq ID No: 380 & 381	103768	AF086009	Hs.296398	gb:Homo sapiens full length insert cDNA
	Seq ID No: 382 & 383	417512	X76534	Hs.82226	glycoprotein (transmembrane) nmb
	Seq ID No: 384 & 385	425266	J00077	Hs.155421	alpha-fetoprotein
	Seq ID No: 386 & 387	424503	NM_002205	Hs.149609	Integrin, alpha 5 (fibronectin receptor,
50	Seq ID No: 388 & 389	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	Seq ID No: 390 & 391	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	Seq ID No: 392 & 393	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	Seq ID No: 394 & 395	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	Seq ID No: 396 & 397	415138	C18356	Hs.285944	tissue factor pathway inhibitor 2
55	Seq ID No: 398 & 399	418506	AA084248	Hs.85339	G protein-coupled receptor 39
	Seq ID No: 400 & 401	423961	D13666	Hs.136348	periostin (OSF-2os)
	Seq ID No: 402 & 403	414812	X72755	Hs.77367	monokine induced by gamma interferon
	Seq ID No: 404 & 405	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
	Seq ID No: 406 & 407	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
60	Seq ID No: 408 & 409	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse
	Seq ID No: 410 & 411	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy
	Seq ID No: 412 & 413	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypothe
	Seq ID No: 414 & 415	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated
	Seq ID No: 416 & 417	411789	AF245505	Hs.72157	Adiclan
65	Seq ID No: 418 & 419	428698	AA852773	Hs.334838	KIAA1866 protein
	Seq ID No: 420 & 421	450098	W27249	Hs.8109	hypothetical protein FLJ21080
	Seq ID No: 422 & 423	421552	AF026692	Hs.105700	secreted frizzled-related protein 4
	Seq ID No: 424 & 425	452747	BE153855	Hs.61460	lg superfamily receptor LNIR
	Seq ID No: 426 & 427	450375	AA009647		a disintegrin and metalloproteinase doma
	Seq ID No: 428 & 429	426215	AW963419	Hs.155223	stannocalcin 2
70	Seq ID No: 430 & 431	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin
	Seq ID No: 432 & 433	432201	AJ538613	Hs.298241	Transmembrane protease, serine 3
	Seq ID No: 434 & 435	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph
	Seq ID No: 436 & 437	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447
	Seq ID No: 438 & 439	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn
75	Seq ID No: 440 & 441	447033	AJ357412	Hs.157601	ESTs
	Seq ID No: 442 & 443	447033	AJ357412	Hs.157601	ESTs
	Seq ID No: 444 & 445	447033	AJ357412	Hs.157601	ESTs
	Seq ID No: 446 & 447	115522	BE614387	Hs.333893	c-Myc target JPO1
	Seq ID No: 448 & 449	410418	D31382	Hs.63325	transmembrane protease, serine 4
80	Seq ID No: 450 & 451	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
	Seq ID No: 452 & 453	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
	Seq ID No: 454 & 455	452461	N78223	Hs.108106	transcription factor
	Seq ID No: 456 & 457	412420	AL035668	Hs.73853	bone morphogenetic protein 2
	Seq ID No: 458 & 459	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara
85	Seq ID No: 460 & 461	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon

Seq ID No: 462 & 463	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365012.1 [H.ssa
Seq ID No: 464 & 465	402075			ENSP00000251056: Plasma membrane calcium
Seq ID No: 466 & 467	421110	AJ250717	Hs.1355	cathepsin E
Seq ID No: 468 & 469	451668	Z43948	Hs.326444	cartilage acidic protein 1
Seq ID No: 470 & 471	451668	Z43948	Hs.326444	cartilage acidic protein 1
Seq ID No: 472 & 473	451668	Z43948	Hs.326444	cartilage acidic protein 1
Seq ID No: 474 & 475	422282	AF019225	Hs.114309	apolipoprotein L
Seq ID No: 476 & 477	425852	AK001504	Hs.159551	death receptor 6, TNF superfamily member
Seq ID No: 478 & 479	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),
Seq ID No: 480 & 481	427747	AW411425	Hs.180655	serine/threonine kinase 12
Seq ID No: 482 & 483	420281	AI623693	Hs.323494	Predicted cation efflux pump
Seq ID No: 484 & 485	405932			C15000305:gij3806122[gb]AAC69198.1 [AF0
Seq ID No: 486 & 487	405932			C15000305:gij3806122[gb]AAC69198.1 [AF0
Seq ID No: 488 & 489	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane
Seq ID No: 490 & 491	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (C/
Seq ID No: 492 & 493	417079	U65590	Hs.81134	interleukin 1 receptor antagonist
Seq ID No: 494 & 495	430890	X54232	Hs.2699	glypican 1
Seq ID No: 496 & 497	419721	NM_001650	Hs.288650	aquaporin 4
Seq ID No: 498 & 499	444471	AB020684	Hs.11217	KIAA0877 protein
Seq ID No: 500 & 501	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote
Seq ID No: 502 & 503	433800	AI034351	Hs.135150	lung type-I cell membrane-associated gly
Seq ID No: 504 & 505	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
Seq ID No: 506 & 507	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
Seq ID No: 508 & 509	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte
Seq ID No: 510 & 511	410407	X65839	Hs.63287	carbonic anhydrase IX
Seq ID No: 512 & 513	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s
Seq ID No: 514 & 515	412719	AW016610	Hs.816	ESTs
Seq ID No: 516 & 517	417034	NM_006183	Hs.80962	neurotensin
Seq ID No: 518 & 519	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
Seq ID No: 520 & 521	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kalini
Seq ID No: 522 & 523	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
Seq ID No: 524 & 525	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
Seq ID No: 526 & 527	418663	AK001100	Hs.41690	desmocollin 3
Seq ID No: 528 & 529	418663	AK001100	Hs.41690	desmocollin 3
Seq ID No: 530 & 531	429610	AB024937	Hs.211092	LUNX protein: PLUNC (palate lung and nas
Seq ID No: 532 & 533	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
Seq ID No: 534 & 535	431846	BE019924	Hs.271580	uropod 1B
Seq ID No: 536 & 537	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
Seq ID No: 538 & 539	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
Seq ID No: 540 & 541	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
Seq ID No: 542 & 543	428484	AF104032	Hs.184501	solute carrier family 7 (cationic amino
Seq ID No: 544 & 545	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
Seq ID No: 546 & 547	417389	BE260964	Hs.82045	midline (neurite growth-promoting factor
Seq ID No: 548 & 549	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
Seq ID No: 550 & 551	417542	J04129	Hs.82269	progesterone-associated endometrial prote
Seq ID No: 552 & 553	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
Seq ID No: 554 & 555	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
Seq ID No: 556 & 557	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
Seq ID No: 558 & 559	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B
Seq ID No: 560 & 561	418462	BE001596	Hs.85266	integrin, beta 4
Seq ID No: 562 & 563	410274	AA381807	Hs.61782	hypoxia-inducible protein 2
Seq ID No: 564 & 565	439606	W79123	Hs.58561	G protein-coupled receptor 87
Seq ID No: 566 & 567	404877			NM_005365:Homo sapiens melanoma antigen,
Seq ID No: 568 & 569	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
Seq ID No: 570 & 571	418543	NM_005329	Hs.85962	hyaluronan synthase 3
Seq ID No: 572 & 573	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
Seq ID No: 574 & 575	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
Seq ID No: 576 & 577	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
Seq ID No: 578 & 579	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
Seq ID No: 580 & 581	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
Seq ID No: 582 & 583	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
Seq ID No: 584 & 585	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
Seq ID No: 586 & 587	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
Seq ID No: 588 & 589	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
Seq ID No: 590 & 591	409420	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kalini
Seq ID No: 592 & 593	332180	AF134160	Hs.7327	claudin 1
Seq ID No: 594 & 595	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
Seq ID No: 596 & 597	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
Seq ID No: 598 & 599	439223	AW238299	Hs.250618	UL16 binding protein 2
Seq ID No: 600 & 601	409757	NM_001898	Hs.123114	cystatin SN
Seq ID No: 602 & 603	428969	AF120274	Hs.194689	artemin
Seq ID No: 604 & 605	428969	AF120274	Hs.194689	artemin
Seq ID No: 606 & 607	428969	AF120274	Hs.194689	artemin
Seq ID No: 608 & 609	428969	AF120274	Hs.194689	artemin
Seq ID No: 610 & 611	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
Seq ID No: 612 & 613	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
Seq ID No: 614 & 615	414774	X02419	Hs.77274	plasminogen activator, urokinase
Seq ID No: 616 & 617	407944	R34008	Hs.239727	desmocollin 2
Seq ID No: 618 & 619	407944	R34008	Hs.239727	desmocollin 2
Seq ID No: 620 & 621	457489	AI693815	Hs.127179	cryptic gene
Seq ID No: 622 & 623	429547	AW009166	Hs.99376	ESTs
Seq ID No: 624 & 625	407242	M18728		gb:Human nonspecific crossreacting anti
Seq ID No: 626 & 627	407242	M18728		gb:Human nonspecific crossreacting anti
Seq ID No: 628 & 629	407242	M18728		gb:Human nonspecific crossreacting anti
Seq ID No: 630 & 631	444006	BE395085	Hs.10086	type 1 transmembrane protein Fn14

	Seq ID No: 632 & 633	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase domain
	Seq ID No: 634 & 635	422109	S73265	Hs.1473	gastrin-releasing peptide
	Seq ID No: 636 & 637	419235	AW470411	Hs.288433	neurotrophin
5	Seq ID No: 638 & 639	449048	Z45051	Hs.22820	similar to S68401 (cattle) glucose induc
	Seq ID No: 640 & 641	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy
	Seq ID No: 642 & 643	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu
	Seq ID No: 644 & 645	448243	AW369771	Hs.52620	integrin, beta 8
	Seq ID No: 646 & 647	426427	M86699	Hs.169840	TTK protein kinase
	Seq ID No: 648 & 649	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6
10	Seq ID No: 650 & 651	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6
	Seq ID No: 652 & 653	428450	NM_014791	Hs.184339	KIAA0175 gene product
	Seq ID No: 654 & 655	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,
	Seq ID No: 656 & 657	453392	U23752	Hs.32864	SRY (sex determining region Y)-box 11
	Seq ID No: 658 & 659	426514	BE616533	Hs.170195	bone morphogenetic protein 7 (osteogenic
15	Seq ID No: 660 & 661	425776	U25128	Hs.159499	parathyroid hormone receptor 2
	Seq ID No: 662 & 663	425776	U25128	Hs.159499	parathyroid hormone receptor 2
	Seq ID No: 664 & 665	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha
	Seq ID No: 666 & 667	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7
	Seq ID No: 668 & 669	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
20	Seq ID No: 670 & 671	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
	Seq ID No: 672 & 673	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
	Seq ID No: 674 & 675	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
	Seq ID No: 676 & 677	410001	AB041036	Hs.57771	kallikrein 11
25	Seq ID No: 678 & 679	426501	AW043782	Hs.293616	ESTs
	Seq ID No: 680 & 681	408369	R38438	Hs.182575	solute carrier family 15 (H777) transport
	Seq ID No: 682 & 683	445413	AA151342	Hs.12677	CGI-147 protein
	Seq ID No: 684 & 685	422424	AJ186431	Hs.296638	prostate differentiation factor
	Seq ID No: 686 & 687	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,
30	Seq ID No: 688 & 689	420610	AJ683183	Hs.99348	distal-less homeo box 5

TABLE 15B

	Pkey:	Unique Eos probeset identifier number	
35	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
	Pkey	CAT Number	Accession
	309931	AW341683	
40	330493	33264_5	M27826 R78416 AA307645 AW957879 AW957800 AA633529 H03662
	439285	47065_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AJ346341 AJ867454 N54784 AI655270 AI421279 AW014882 AA775552 N62351 N59253 AA626243 AJ341407 BE175639 AA456968 AJ358918 AA457077
	450375	83327_1	AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532
			AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AJ161014 AA099554 R69087
45	451320	86576_1	AW118072 AI631982 T15734 AA224195 AJ701458 W20198 F26326 AA890570 N90552 AW071907 AJ671352 AJ375892 T03517 R88265 AI124088 AA224388 AJ084316 AJ354686 T33652 AJ140719 AJ720211 T03490 AJ372637 T15415 AW205836 AA630384 T03515 T33230 AA017131 AA443303 T33623 AJ222556 T33511 T33785 AJ419606 D55612

TABLE 15C

50

Pkey:

Unique number corresponding to an Eos probeset

Ref:

Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.

Strand:

Indicates DNA strand from which exons were predicted.

55

NL_position:

Indicates nucleotide positions of predicted exons.

60

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
403329	8516120	Plus	96450-96598
403478	9958258	Plus	116458-116564
404440	7528051	Plus	80430-81581
404677	1519284	Plus	1095-2107
405770	2735037	Plus	61057-62075
405932	7767812	Minus	123525-123713

65

Table 16

Seq ID NO: 1 DNA sequence
Nucleic Acid Accession #: NM_001216
Coding sequence: 43..1422

5

	1	11	21	31	41	51	
	GCCCGTACAC	ACCGTGTGCT	GGGACACCCC	ACAGTCAGCC	GCATGGCTCC	CCTGTGCCCC	60
10	AGCCCCCTGGC	TCCCTCTGTT	GATCCCGGCC	CCTGCTCCAG	GCCTCACTGT	GCAACTGCTG	120
	CTGTCACTGC	TGCTTCTGAT	GCTGTCCAT	CCCCAGAGGT	TGCCCCGGAT	GCAGGAGGAT	180
	TCCCCCTTGG	GAGGAGGCTC	TTCTGGGGAA	GATGACCCAC	TGGGCGAGGA	GGATCTGCCC	240
	AGTGAAGAGG	ATTCACCACG	AGAGGAGGAT	CCACCCGGAG	AGGAGGATCT	ACCTGGAGAG	300
	GAGGATCTAC	CTGGAGAGGA	GGATCTACCT	GAAGTTAAGC	CTAAATCAGA	AGAAGAGGGC	360
15	TCCCTGAAGT	TAGAGGATCT	ACCTACTGTT	GAGGCTCCTG	GAGATCTCTA	AGAACCCTAG	420
	AATAATGCCC	ACAGGAGCAA	AGAAGGGGAT	GACCAGAGTC	ATTGGCGCTA	TGGAGGGGAC	480
	CGCCCTTGGC	CCCGGGTGTG	CCAGCCCTGC	GCGGGCCGCT	TCCAGTCCCC	GGTGGATATC	540
	CGCCCCCAGC	TCCCGCCCTT	CTGCCCGGCC	CTGCCCGCCC	TGGAATCTCT	GGGCTTCCAG	600
	CTCCCGCGGC	TCCCAAGACT	GCGCCTGCGC	AACAATGGCC	ACAGTGTGCA	ACTGACCTTG	660
20	CCTCCTGGGC	TAGAGATGGC	TCTGGGTCCC	GGGCGGGAGT	ACCGGCTCTT	GCAGCTGCAT	720
	CTGCACCTGG	GGGCTGCAGG	TGCTCCGGGC	TGGGAGCACA	CTGTGGAAGG	CCACCGTTTC	780
	CCTGCGGAGA	TCCAGCTGGT	TCACTCAGC	ACCGCCTTTG	CCAGAGTTGA	CGAGGCGCTG	840
	GGGCGCCCGG	GAGGCTTGGC	CGTGTGGGCC	GCCTTCTGCG	AGGAGGGCCC	GGAGAAAAC	900
	AGTGCCCTATG	AGCAGTTGCT	GTCTCGCTTG	GAAGAAATCG	CTGAGGAAGG	CTCAGAGACT	960
25	CAGGTCCAG	GACTGGACAT	ATCTGCACTC	CTGCCCTCTG	ACTTCAGCCG	CTACTTCCAA	1020
	TATGAGGGGT	CTCTGACTAC	ACCGCCCTGT	GCCCAGGGTG	TCATCTGGAC	TGTGTTTAAC	1080
	CAGACAGTGA	TGCTGAGTGC	TAAGCAGCTC	CACACCCCTC	CTGACACCCT	GTGGGGACCT	1140
	GGTGACTCTC	GGCTACAGCT	GAACCTCCGA	GCGACGCAAG	CTTTGAATGG	GCGAGTGATT	1200
	GAGGCTCTCT	TCCCTGCTGG	AGTGGACAGC	AGTCTCTCGG	CTGCTGAGCC	AGTCCAGCTG	1260
30	AATTCTGTGC	TGGCTGTCTG	TGACATCTTA	GCCCTGGTTT	TGGCCCTCTT	TTTGTCTGTC	1320
	ACCAGCGTGC	CGTCTCTTGT	GCAGATGAGA	AGGCAGCACA	GAAGGGGAAC	CAAGGGGGGT	1380
	GTGAGCTACC	GCCACGACGA	GGTAGCCGAG	ACTGGAGCCT	AGAGGCTGGA	TCTTGGAGAA	1440
	TGTGAGAAGC	CAGCCAGAGG	CATCTGAGGG	GGAGCCGGTA	ACTGTCTCTG	CCTGCTCATT	1500
35	ATGCCACTTC	CTTTTAACCTG	CCAAGAAATT	TTTTAAATA	AATATTATA	AT	

Seq ID NO: 2 Protein sequence:
Protein Accession #: NP_001207

	1	11	21	31	41	51	
	MAPLCPSPL	PLIPAPAPG	LTVQLLSLL	LLMPVHPQRL	PRMQEDSPLG	GGSSGEDDPL	60
	GEEDLPSEED	SPREEDPPGE	EDLPGEEDLP	GEEDLPVHP	KSEBEGSLKL	EDLPTEVAPG	120
	DPQEPQNAH	RDKEGDDQSH	WRYGGDPFPW	RVSPACAGRF	QSPVDIRPOL	AAFCPALRPL	180
45	ELLGFLPL	PELRLRNNGH	SVQLTLPPGL	EMALGPGREY	RALQLHLHWG	AAGRPGSEHT	240
	VEGHRFPFPI	HVVELSTAFI	RVDEALGRPG	GLAVLAAPLE	EGPEENSAYE	QLLSRLERIA	300
	EEGSETQVPG	LDISALLPSD	FSRYFQYEGS	LTPPPCAQGV	IWTVENQTVM	LSAKQLHTLS	360
	DTLWGPDSR	LQLNFRATQP	LNGRVIEASF	PAGVDSFRA	AEPVQLNSCL	AAGDILALVF	420
	GLLFAVTSVA	FLVQMRQRHR	RGTKGGVSYS	PAEVAETGA			

50

Seq ID NO: 3 DNA sequence
Nucleic Acid Accession #: BC013923
Coding sequence: 438-1391

	1	11	21	31	41	51	
	AGCGGGGTTG	TCTATTAAC	TGTTCAAAAA	GTATCAGGAG	TTGTCAAGGC	AGAGAAGAGA	60
	GTGTTTGCAA	AAGGGGGAAA	GTAGTTTGCT	GCCTCTTTAA	GACTAGGACT	GAGAGAAGA	120
	AGAGGAGAGA	GAAAGAAAGG	GAGAGAAGTT	TGAGCCCCAG	GCTTAAGCCT	TTCCAAAAAA	180
	TAATAATAAC	AATCATCGCG	GGCGGCAGGA	TGCGCCAGAG	GAGGAGGGAA	CGCTTTTTTT	240
60	TGATCTGAT	TCCAGTTTGC	CTCTCTCTTT	TTTTCCCCCA	AATTATCTTT	CGCTGATTTT	300
	TCTCTCGCGA	GCCCTCGCCT	CCGACACCCC	CGGCGCGCCT	CCCTCTCTCC	TCTCCCCCGG	360
	CCCGCGGGCC	CCCCAAGTCT	CCGCGCGGCG	CGAGGCTCGG	CGGCGCGCGG	CGGCGCGGGC	420
	CCGCGCACAG	CGCCCGCATG	TACAACATGA	TGGAGACGGA	GCTGAAGCCG	CGGCGCCCGC	480
	AGCAAACTTC	GGGGGGCGCG	GGCGGCAACT	CCACCGCGCG	GGCGCGCGCG	GGCAACAGAA	540
65	AAAAACAGCC	GGACCGCGTC	AAGCGGCCCA	TGAATGCCCT	CATGGTGTGG	TCCCGCGGGC	600
	AGCGCGCAAA	GATGGCCAG	GAGAAACCCA	AGATGCACAA	CTCGGAGATC	AGCAAGCGCC	660
	TGGGCGCCGA	GTGGAACCTT	TTGTCCGAGA	CGGAGAAGCG	GCCGTTTATC	GACGAGGCTA	720
	AGCGGCTGCG	AGCGCTGCAC	ATGAAGGAGC	ACCCGAGTTA	TAAATACCGG	CCCCGCGGGA	780
	AAACCAAGAG	GCTCATGAAG	AAGGATAAGT	ACACGCTGCC	CGGCGGGCTG	CTGGCCCCCG	840
70	GCGGCAATAG	CATGGCGAGC	GGGGTGGGGG	TGGGCGCGCG	CCTGGGCGCG	GGCGTGAAAC	900
	AGCGCATGGA	CAGTTACGCG	CACATGAACG	GCTGGAGCAA	CGGCAGCTAC	AGCATGATGC	960
	AGGACCAGCT	GGGCTACCCG	CAGCACCCCG	GCCTCAATGC	GCACGGCGCA	GCGCAGATGC	1020
	AGCCCATGCA	CCGCTACGAC	GTGAGCGCCC	TGCAGTACAA	CTCCATGACC	AGCTCGCAGA	1080
	CCTACATGAA	CGGCTCGCCC	ACCTACAGCA	TGCTCTACTC	GCAGCAGGGC	ACCCCTGGCA	1140
75	TGGCTCTTGG	CTCCATGGGT	TGCGTGGTCA	AGTCCGAGGC	CAGCTCCAGC	CCCCCTGTGG	1200
	TTACCTCTTC	CTCCCACTCC	AGGGCGCCCT	GCCAGGCGCG	GGACCTCCCG	GACATGATCA	1260
	GCATGTATCT	CCCCGGCGCC	GAGGTGCGCG	AACCGCGCGC	CCCCAGCAGA	CTTCACATGT	1320
	CCACAGACTA	CCAGAGCGCG	CCGGTGCCCG	GCACGGCCAT	TAAACGACAC	CTGCCCCCTC	1380
	CACACATGTG	AGGGCGCGAC	AGCGAACTGG	AGGGGGGAGA	AATTTTCAAA	GAAAAACGAG	1440
80	GGAAATGGGA	GGGGTGCAAA	AGAGGAGAGT	AAGAAACAGC	ATGGAGAAAA	CCCGGTACGC	1500
	TCAAAAAAAA	AAAAAAAATA	AAAAATCCAT	CACCCACAGC	AAATGACAGC	TGCAAAAGAG	1560
	AACACCAATC	CCATCCACAC	TCACGCAAAA	ACCGGATGCG	CGACAAGAAA	ACTTTTATGA	1620
	GAGAGATCTT	GGACTTCTTT	TKGGGGGACT	ATTTTGTGAC	AGAGAAAAAC	TGGGGAGGGT	1680
	GGGGAGGGCG	GGGGAATGGA	CCTTGTATAG	ATCTGGAGGA	AAGAAAGCTA	CGAAAAACTT	1740
85	TTTAAAGTTT	CTAGTGGTAT	GGTAGGAGCT	TTGCAGGAAG	TTTGCAGGAAG	TCTTTACCAA	1800
	TAATATTTAG	AGCTAGTCTC	CAAGCGACGA	AAAAAATGTT	TTAATATTGT	CAAGCAACTT	1860
	TTGTACAGTA	TTTATCGAGA	TAAACATGGC	AATCAAAATG	TCATTGTTT	ATAAGCTGAG	1920

AATTGGCCAA TATTTTTCAA GGAGAGGCTT CTGCTGAAAT TTTGATTCTG CAGCTGAAAT 1980
TTAGGACAGT TGCACACGTT AAAAGAAGAA AATTATTCAA ATTTGGACAT TTTAATTGTT 2040
TAAAAATGT ACAAAAGGAA AAAATTAGAA TAAGTACTGG CGAACCATCT CTGTGGTCTT 2100
5 GTTTAAAAAG GCACAAAGTT TTAGACTGTA CTTAAATTTA TAACCTACTG TTTAAAGCAA 2160
AAATGGCCAT GCAGTTGAC ACGTGTGTA ATTATAATA GCTTTGTTC GATCCCACT 2220
TTCCATTTT TTCAATAAA AAAAACCATG AAATTACTGT GTTTGAAATA TTTTCTTATG 2280
GTTTGTAAAT TTTCTGTAAA TTTATTGTGA TATTTTAAGG TTTTCCCCC TTTATTTTCC 2340
GTAGTTGTAT TTTAAAGAT TCGGCTCTGT ATTATTGAA TCAGTCTGCC GAGAAATCCAT 2400
10 GTATATAATT GAACATAAT CATCCTTATA ACAGGTACAT TTTCAACTTA AGTTTCTACT 2460
CCATTATGCA CAGTTTGAGA TAAATAAAT TTTGAAATAT GGACACTGAA AAAAAAATA 2520
AAAAAACAA AACAAAAAA CAAAAACAA AACAGAAAA AACAAAAAA AAAAAAATA 2580
CACACACAA AACAAAAAA AAAAAAAGA AACAAACAA CAACACAAA CAACACAAA 2640
CCACACACA AACAAACA CACAGAGGG

Seq ID NO: 4 Protein sequence:
Protein Accession #:CAA83435.1

1 11 21 31 41 51
MYNMMETELK PPGPQOTSGG GGNSTAAAA GGNQKNSPDR VKRPMNAPMV WSRGQRRKMA 60
QENPKMHNSE ISKRLGAEWK LLSETEKRPF IDEAKRLRAL HMKHPDYKY RPRRKTKTLM 120
KDKYTLPGG LLAPGGNSMA SGVGVGAGLG AGVNQRMSY AEMNGWSNGS YSMQDQLGY 180
PQHPGLNAHG AAQMPMERY DVSAQYNMS TSSQTYMNGS PTYSMSYSQQ GTPGMALGSM 240
25 GSVVKSEASS SPPVVTSSSH SRAPQAGDL RDMISMYLPG AEVPEPAAPS RLHMSQHYQS 300
GPVPGTAING TLPLSHM

Seq ID NO: 5 DNA sequence
Nucleic Acid Accession #: U91618
Coding sequence: 29-541

1 11 21 31 41 51
CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60
CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAAATGAA 120
35 AGCATTAGAA GCAGATTTC TTACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180
TCCCTCTGGG AAGATGACTC TGCTAAATGT TTGCACTCTT GTAAATAATT TGAACAGCCC 240
AGCTGAGGAA ACAGGAGAAG TTCATGAAGA GGAGCTTGTT GCAAGAAGGA AACTTCTTAC 300
TGCTTTAGAT GGCTTTAGCT TGAAGCAAT GTTGACAATA TACCAGCTCC ACAAAATCTG 360
40 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA 420
TGACAAAAAT GCGAAGGAAG AAGTCATAAA GAGAAAAATT CCTTATATTC TGAACGGCA 480
GCTGTATGAG AATAACCCA GAAGACCCCTA CATACTCAAA AGAGATTCTT ACTATTACTG 540
AGAGAAATAA TCATTATTAT CATGTGATT GTGATTCATC ATCCCTTAAT TAAATATCAA 600
ATTATATTTG TGTGAAAATG TGACAAACAC ACTTATCTGT CTCCTTCTCA ATTGTGGTTT 660
45 ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAAA TAAATCTAAA 720
TCTTCAAAA AAAAAAATAA AATGGGGCC GCAATT

Seq ID NO: 6 Protein sequence:
Protein Accession #: AAB50564

1 11 21 31 41 51
MMAGMKIQLV CMLLLAPSSW SLCSDSEREM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60
55 VCSLVNINLS PAEETGEVHE EELVARRKLP TALDGFSLA MLTIYQLHKI CHSRAPQWHE 120
LIQEDILDTG NDRNGKEEVI KRKIPYILKR QLVENKPRRP YILKRDYIYY

Seq ID NO: 7 DNA sequence
Nucleic Acid Accession #: NM_006536.2
Coding sequence: 109-2940

1 11 21 31 41 51
ACCTAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60
65 ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120
AGCATTGCAG GTCTATTTTG CAACCTGAAG TTGTGACTC TCTGTGTTGC CTTAAGTTCA 180
GAATCCCATC TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC 240
ATTGCAATTA ATCTCAGGT ACCTGAGAAT CAGAACCTCA TCTCAACAT TAAGGAAATG 300
70 ATAACCTGAAG CTTCATTTTA CCTATTAAAT GCTACCAAGA GAAGAGTATT TTTGAGAAA 360
ATAAAGATT TAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAAGAA 420
TCATATGAAA AGGCAAAATG CATAGTGAAT GACTGTGATG GGGCAGATGG AGATGATCCA 480
TACACCTAC AATACAGAGG GTGTGGAAA GAGGGAATAT ACATTCATT CACACCTAAT 540
75 TTCTACTGA ATGATAACTT AACAGCTGGC TACGATCAC GAGGCGAGT GTTTGTCCAT 600
GAATGGGCC ACCTCCGTTG GGGTGTGTTG GATGAGTATA ACAATGACAA ACCTTCTTAC 660
ATAAATGGGC AAAATCAAAT TAAAGTGACA AGGTGTTTAT CTGACATCAC AGGCATTTT 720
GTGTGTGAAA AAGGTCTTGG CCCCAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780
GGATGCACCT TTATCTACAA TAGCACCAGG AATGCAACTG CATCAATAAT GTTCATGCAA 840
80 AGTTTATCTT CTGTGGTTGA ATTTTGAAT GCAAGTACCC ACAACCAAGA AGCACCAGAA 900
CTACAGAAC AGATGTGAGC CCTCAGAGT GCATGGGATG TAATCAAGA CTCTGTCTAC 960
TTTCAACCA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCCTCCAC ATTCTGCTT 1020
GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTGCTGGATG TGTCCAGCAA GATGGCAGAG 1080
GCTGACAGC TCCTTCAACT ACAACAAGCC GCAGAAATTT ATTTGATGCA GATTGTGAA 1140
ATTATACCT TCGTGGGCAT TGCCAGTTTC GACAGCAAAG GAGAGATCAG AGCCAGCTA 1200
85 CACCAATTA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260
TCAGCTAAA GACACATCAG CATTGTGTTA GGGCTTAAGA AAGGATTGA GGTGGTTGAA 1320
AAACTGAATG GAAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG 1380
CTTCTTGCCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAAATCA CTCCTATGCC 1440

CTGGGTTTCAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCCTACAGG AGGTTTAAAG 1500
TCTTTTGTTC CAGATATATC AAATCCCAAT AGCATGATTG ATGCTTTTCAG TAGAATTTC 1560
TCTGGAACCTG GAGACATTTT CCAGCAACAT ATTACAGCTTG AAAGTACAGG TGAAAAATGTC 1620
AAACCTCAC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680
5 ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCCTGAGA TTATATTATT TGATCCTGAT 1740
GGACGAAAAA ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800
TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCCTGAACAA TACCCATCAT 1860
TCTCTGCAAG CCTGAAAGT GACAGTGACC TCTCGGCGCT CCAACTCAGC TGTGCCCCCA 1920
GCCACTGTGG AAGCCTTTGT GGAAGAGAG AGCCTCCATT TTCTCATOC TGTGATGATT 1980
10 TATGCCAATG TGAACAGGG ATTTTATCCC ATTCTTAATG CCAGTGTGAC TGCCACAGTT 2040
GAGCCAGAGA CTGGAGATCC TGTTCAGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100
GTTATAAAAA ATGATGGAAT TTAATCGAGG TATTTTTTCT CCTTTGCTGC AAATGGTAGA 2160
TATAGCTTGA AAGTGCATGT CAATCACTCT CCCAGCATAA GCACCCCAAC CCACTCTATT 2220
CCAGGGAGTC ATGCTATGTA TGTACCAGGT TACACAGCAA ACGTAAATAT TCAGATGAAT 2280
15 GCTCCAAGGA AATCAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340
AGCTCAGAGG GCTCCTTTTC AGTGCTGGGA GTTCCAGCTG GCCCCACACC TGATGTGTTT 2400
CAACCATAGA AAATTATTGA CCTGGAAGCT GTAAAAGTAG AAGAGGAATT GACCCTATCT 2460
TGGACAGCAC CTGGAGAAGA CTTTGATCAG GGCCAGGCTA CAAGCTATGA AATAAGAATG 2520
AGTAAAGTC TACAGAATAT CCAAGATGAC TTTAACAATG CTATTTTAGT AAATACATCA 2580
20 AAGCGAATC GCAGCAAGC TGGCATCAGG GAGATATTGA CGTTCTCACC CCAGATTTC 2640
ACGAATGAGC CTGAACATCA GCCAATGGA GAACACATG AAAGCCACAG AATTATGTT 2700
GCAATACAGG CAATCAGTAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCACGGG 2760
CCTCTGTGTTA TTCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820
25 GGAQTITTAAG CAGCAATGGG TTTGATAGGA ATCATTTGCC TTATTATAGT TGTGACACAT 2880
CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATAA 2940
ATAAATATCC AAAGTGTCTT CCTTCTTGA TATAAGACCC ATGGCCCTCG ACTACAAAAA 3000
CATACTAACA AAGTCAAAAT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060
ATACAGATAA GATTTTACA TGGTAGATCA ACAATTCTTT TTGGGGGTAG ATTAGAAAAC 3120
30 CCTTACACTT TGGCTATGAA CAAATAATAA AAATTATTCT TAAAGTAAT GTCTTTAAAG 3180
GCAAGGATAA GGTAAAGTC GGACCAAGTG CAAGGAAAGT TTGTTTTATT GAGTGGGAAA 3240
AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA 3300
TCATTGATT ACTTTGATT ATTTTCTTT TCTCCTTATC TGTGAGTAC AGGTGTCTG 3360
35 TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAG CTTCTTACCT 3420
CTTGCTATT TGTATTATAT ATTTACAGAT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480
TTTCACTGTA AGAGGTAACT TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540
TTTATGACAA AGGTCTATTG AATTATTATG TMTGTAAGTT TCTACTCCCA TCAAAGCAGC 3600
TTTCTAAGTT TATTGCCITG GGTATTATG GAATGATAGT TATAGCCCN TATAATGCCT 3660
TACCTAGGAA A

Seq ID NO: 8 Protein sequence:
Protein Accession #: NP_006527.1

1 11 21 31 41 51
45 MTQRSIAGPI CNLKFVTLIV ALSSELPLFG AGVQLQDNGY NGLLIAINPQ VPENQNLIEN 60
IKEMITEASF YLENATKRRV FFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAH 120
GDDPYTLQYR GCKGEGKVIH FTFNFIINDN LTAGYGSRRG VFVREMAHLR WGVFDEYNND 180
KPFYINGNQK IKVTRCSSDI TGIFVCEKSP CPQENCIISK LFKEGCTFIY NSTQNATASI 240
50 MFMQSLSSVV EFCNASTHNQ EAPNLQNMCM SLRSANDVIT DSADPHHSFP MNGTELPFP 300
TFSLVQAGDK VVCLVLDVSS KMAEADRLLO LQQAEEFYLM QIVEIHTFVG IASFDSKGEI 360
RAQLHQINSN DRKLLVSYL PTTVSAKTDI SICSGLKKGF EVVEKLNGKA YGSMVLVTS 420
GDDKLLGNDL PTVLSSGSTI HSIALGSSAA PNLEELSRIT GGLKPFVFDI SNNSMSIDAF 480
SRISSGTGDI PQHIQLEST GENVKPHEQL KNTVTVDNVT GNDTMFLVTV QASGPPEIIL 540
55 FDPDGRKYIT NNFITNLFR TASLWIPGTA KPGHWITLIN MTHSLQALK VVTVSRASNS 600
AVPPATVEAF VERDSLHFFH PVMIIYANVKQ GPYFILNATV TATVEPETGD FVTLRLDDG 660
AGADWIKNDG IYSRYFFSFA ANGRYSKLVH VNRSPSISTP AHSIPGSHAM YVPGYTANGN 720
IQMNAPRKSV GRNEERKRWG FSRVSSGGSP SVLGVPAAPH PDVPPCKII DLEAVKVEEE 780
LTLSTWAPGE DFDQQAQTSY EIRMSKSLQN IQDDFNAIL VNTSKRNPQ AGIREIPTFS 840
60 PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900
LILKGVLTAM GLIGIICLII VVTHEILSRK KRADKKENGK KLL

Seq ID NO: 9 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 336-632

1 11 21 31 41 51
70 CTCCTCTCAC CCGGTCCAG GATGCCAGT CCCCAGACA CCTCCACTT CCCACTGTGG 60
CCTGGGTGGG CTCAGGGGCT GCCCTTGACC TGGCTAGAG CCCTCCCCCA GCTGGTGGTG 120
GAGCTGGCAC TCTCTGGGAG GGAGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180
CAGGGTTTGG TGGGATCAGG TTGAGGCAGG TTTGGTTTCC TTAATAATGCC AAGTTGGGGG 240
CCAGTGGGGC CCACATATAA ATCTCACC CCAGGAGCCTG GCTGCCCTGC TCTCCTTCT 300
75 GGGTCTGTCT CTGCCACCTG GTCTGCCACA GATCCATGAT GTGCAGTCT CTGGAGCAGG 360
CGCTGGCTGT GCTGGTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420
AGCTGAGTAA GGGGGAATG AAGGAACCTC TGCACAAGGA GCTGCCAGC TTTGTGGGGG 480
AGAAAGTGA TGAAGAGGGG CTGAAGAAGC TGATGGGCAG CCTGGATGAG AACAGTGACC 540
AGCAGGTGGA CTTCCAGGAG TATGCTGTT TCTGGCACT CATCACTGTC ATGTGCAATG 600
80 ACTTCTTCCA GGGCTGCCA GACCGACCT GAAGCAGAAC TCTTGACTTC CTGCCATGGA 660
TCTCTTGGGC CCAGGACTGT TGTGCTCTT GAGTTTTGTA TTAATAAAC TTTTTTGTG 720
TGTGTATAAT ATTTTAATG CTCAGTGATG TTCCATAACC CGGCTGGCTC AGCTGGAGTG 780
CTGGGAGATG AGGGCTCCT GATCTCTGCT CCTCTCTGG CTCTGACTCT CCTGGAATC 840
TCTCAAGGC CAGAGCTATG CTTAGGTCT CAATTTTGA ATTTCAACA CCAGCAAAA 900
85 ATTGGAATC GAGATAGGTT GCTGACTTTT ATTTGTGCAA ATAAAGATAT TAAAAAGGC 960
AAATACCA

Seq ID NO: 10 Protein sequence:

Protein Accession #: NP_005969.1

1 11 21 31 41 51
5 MMCSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGEK VDEEGLKKLM 60
GSLDENSQQ VDFQYAVFL ALITVMCNDF FQGCPRDP

10 Seq ID NO: 11 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 336-626

1 11 21 31 41 51
15 CTCCCCAC CCGGCTCCAG GATGCCAGT CCCCACGACA CCTCCCACTT CCCACTGTGG 60
CCTGGGTGGG CTGAGGGGCT GCGCTTGACC TGGCCTAGAG CCCTCCGCCA GCTGGTGGTG 120
GAGCTGGCAC TCTCTGGGAG GAGGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180
CAGGGTTTGG TGGGATCAGG TTGAGGCAGG TTTGGTTTCC TTAATAATGCC AAGTTGGGGG 240
CCAGTGGGCG CCACATATAA ATCCTCACCC TGGGAGCCTG GCTGCCTTGC TCTCCTTCT 300
20 GGGTCTGTCT CTGCCACCTG GTCTGCCACA GATCCATGAT GTGCAGTTCT CTGGAGCAGG 360
CGCTGGCTGT GCTGGTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420
AGCTGAGTAA GGGGGAATG AAGGAACCTC TGCAACAAGA GCTGCCAGC TTTGTGGGGC 480
ATTCCAGAGA ACCATGTGCT GTGAGGGCCT TCCGAGTCCA TCTGTTTAACT CCTGTCAATG 540
GAGACTTGAG AAACAGAGC CCAGAAGGGA AAGTGATTG TCCCAAGATC ACACAGCACT 600
25 GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC AGCCTGGATG AGAAGCTGA 660
CCAGCAGGTG GACTTCCAGG AGTATGCTGT TTTCTGGGCA CTCATCACTG TCATGTGCAA 720
TGACTTCTTC CAGGGCTGCC CAGACCGACC CTGAAGCAGA ACTCTTGACT TCCTGCGCATG 780
GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG TATTCAATAA ACTTTTITTT 840
30 TCTGTTGATA ATATTTTAAAT TGCTCAGTGA TGTTCATAA CCGGCTGGC TCAGCTGGAG 900
TGCTGGGAGA TGAGGGCCTC CTGGATCCTG CTCCTTCTG GGTCTGACT CTCCTGGAAG 960
TCTCTCCAGG GCCAGAGCTA TGCTTTAGGT CTCAATTTTG GAATTTCAA CACCAGCAA 1020
AAATTGGAAG TCGAGATAGG TTGCTGACTT TTATTTTGTC AAATAAGAT ATTAATAAAG 1080
GCAATACCA

35 Seq ID NO: 12 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
40 MMCSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGHS REPCAVRAFR 60
VHLFNPVIGD LRNQSPEGKS DCPKITQHWR KWMRRG

45 Seq ID NO: 13 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 58-354

1 11 21 31 41 51
50 GTGAGCTCAC CATGTGGGGG TGAGGCTGAG AGAAAAAAG TACACAGCCA CAGATCCATG 60
ATGTGCAAGT CTCTGGAGCA GCGCTGGGCT GTGCTGGTCA CTACCTTCCA CAACTACTCC 120
TGCCAAGAGG GCGACAGATT CAAGCTGAGT AAGGGGGAAA TGAAGGAAGT TCTGCACAAG 180
GAGCTGCCCA GCTTTGTGGG GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC 240
AGCCTGAGTG AGAACAGTGA CCAGCAGGTG GACTTCCAGG AGTATGCTGT TTTCTGGGCA 300
55 CTCATCACTG TCATGTGCAA TGACTTCTTC CAGGGCTGCC CAGACCGACC CTGAAGCAGA 360
ACTCTGACT TCTGCGCATG GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG 420
TATTCAATAA ACTTTTITTT TCTGTTGATA ATATTTTAAAT TGCTCAGTGA TGTTCATAA 480
CCCGCTGGGC TCAGCTGGAG TGCTGGGAGA TGAGGGCCTC CTGGATCCTG CTCCTTCTG 540
60 GGTCTGACT CTCCTGGAAG TCTCTCCAAG GCCAGAGCTA TGCTTTAGGT CTCAATTTTG 600
GAATTTCAA CACCAGCAA AAATTGGAAG TCGAGATAGG TTGCTGACTT TTATTTTGTC 660
AAATAAGAT ATTAATAAAG GCAATACCA

Seq ID NO: 14 Protein sequence:
Protein Accession #: NP_005969.1

1 11 21 31 41 51
65 MMCSLEQAL AVLVTTFHKY SCQEGDKFKL SKGEMKELLH KELPSFVGEK VDEEGLKKLM 60
GSLDENSQQ VDFQYAVFL ALITVMCNDF FQGCPRDP

70 Seq ID NO: 15 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 62-358

1 11 21 31 41 51
75 GGAGGTGTG CCGCTGAGTC ACTGCCTGGG CATCTGGGCC TGGAACTCTG GCCACAGATC 60
CATGATGTGC AGTTCTCTGG AGCAGGCGCT GGCTGTGCTG GTCACTACTT TCCACAAGTA 120
80 CTCTGCCCAA GAGGGCGACA AGTTCAAGCT GAGTAAGGGG GAAATGAAGG AACTTCTGCA 180
CAAGGAGCTG CCAGCTTTTG TGGGGGAGAA AGTGGATGAG GAGGGGCTGA AGAAGCTGAT 240
GGGAGCCTG GATGAGAAAC GTGACCAGCA GGTGGACTTC CAGGAGTATG CTGTTTCTCT 300
GGCACTATC ACTGTGATGT GCAATGACTT CTTCCAGGGC TGCCAGAGCC GACCTGGAAG 360
CAGAATCTTT GACTTCTCTG CATGGATCTC TTGGGCGCAG GACTGTTGAT GCCTTTGAGT 420
85 TTTGATTTCA ATAACTTTT TTTGTCTGTT GATAATATTT TAATTGCTCA GTGATGTTCC 480
ATAACCGGCG TGCTCAGCT GAGATGCTGG GAGATGAGGG CCTCTGGAT CCTGCTCCCT 540
TCTGGGCTCT GACTCTCTG GAAATCTCTC CAAGGCCAGA GCTATGCTTT AGTCTCTCAAT 600
TTTGAATTT CAAACACAG CAAAAAATTG GAAATCGAGA TAGGTTGCTG ACTTTTATTT 660

Seq ID NO: 16 Protein sequence:
Protein Accession #: NP_005969.1

5
1 11 21 31 41 51
MMCSSLEQAL AVLVTTTPHYK SCQEGDKPKL SKGEMKELLH KELPSFVGEK VDEBGLKKLM 60
10
GSLDENSDQQ VDPQYAVFL ALITVMCNDF FQGCDFRP

Seq ID NO: 17 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 939-2372

15
1 11 21 31 41 51
20
AAGACGGATT CTCAGACAAG GCTTGCAAAT GCCCGCAGC CATCATTTAA CTGACCCGCG 60
AGAATAGTTA CGTGTGTGCA CCGGACCCCT CCGGATCGCC TAATTTGTCC CTAGTGAGAC 120
CCCGAGGCTC TGCOCGCGCC TGGCTTCTTC GTAGCTGGAT GCATATCGTG CTCGSGGCAG 180
CGCGGGCGCA GGGCAGCGGT TCGCGCACAC CCTAGCACAC ATGAACACGC GCAAGAGCTG 240
AAGCAAGCAC GGTTCCTCAT TCAAAAAGGG AGACAGCCTC TACCGCGATT GTAGAAGAGA 300
CTGTGTGTG AATTAGGGAC CGGAGGCGGT CGAACGAGG AACGGTTCAT CTTAGAGACT 360
25
AATTTTCTGG AGTTTCTGCC CCTGCTCTGC GTACGCCCTC ACGTCACTTC GCCAGCAGTA 420
GCAGAGGCGG CGGCGCGCGC TCCCGGAATT GGGTTGGAGC AGGAGCGCTG CTGGCTGCTT 480
CGCTCGCGCT CTACGCGCTC AGTCCCGGCG GGTAGCAGGA GCCTGGACCC AGGCGCGCGC 540
GGCGGGCGTG AGGCGCGCGA GCCCGGCGCT GAGGTGCATA CCGGACCCCG ATTGCGCATCT 600
30
AACAAGGAAT CTGCGCCCA GAGAGTCCCG GAGCGCGCGC CGGTGCGTGC CCGGCGCGCC 660
GGGCCATGCA GCGACGCGCG CCGCGGAGCT CCGAGCAGCG GTAGCGCCCC CCTGTAAAGC 720
GGTTCGCTAT GCCCGGGCCA CTGTGAACCC TGCCGCGCTG CCGAACACTC TTCGCTCCGG 780
ACCGAGCTAG CCTCTGATAA GCTGGACTCG GCACGCGCGC AACAGCAGC GAGGAGTTAA 840
GAGAGCGCGA AGCGCAGGGA AGGCGTCCCG GCACGCGTGG GGGAAAGCGG CCGGTGCGAGC 900
35
GCGGGGACAG GCCTCGGGC TGGCACTGGC TGCTAGGGAT GTCTGCTGCT ATAAGGTGGC 960
ATGGACCGCG CATGCGCGCG CTCTGGGCGT TCTGCTGGCT GGTGTGGGGC TTCGAGAGGG 1020
CCGCTTTCCG CTGTCCCGCG TCCGTGCAAT GCAGTGCCTC TCGGATCTGG TGCAGCGACC 1080
CTTCTCTGG CATCGTGGCA TTTCCGAGAT TGGAGCCTAA CAGTGTAGAT CCTGAGAAAC 1140
TCACCGAAAT TTTCTATGCA AACCGAAGAA GGTAGAAAT CATCAACGAA GATGATGTTG 1200
40
AAGCTTATGT GGGACTGAGA AATCTGACAA TTGTGGATTC TGGATTAAAA TTTGTGGCTC 1260
ATAAAGCATT TCTGAAACAC AGCAACCTGC AGCACATCAA TTTTACCCGA AACAACTGA 1320
CGAGTTTGTC TAGGAAACAT TCCGCTCACC TTGACTTGTC TGAACGTATC CTGGTGGGCA 1380
ATCCATTATC ATGCTCTGCT GACATTATGT GGATCAAGAC TCTCCAAGAG GCTAAATCCA 1440
GTCCAGACAC TCAAGATTGT TACTGCTGTA ATGAAAGCAG CAAGAATATT CCGCTGGCAA 1500
45
ACCTGCAGAT TCTCGAATGT GGTITGCCAT CTGCAAACTC GGGCGCACCT AACCTCACTG 1560
TGGAGGAAGG AAGTCTATC ACATTATCCT GTAGTGTGGC AGGTGATCCG GTTCTTAATA 1620
TGATTTGGGA TGTGTGTAAC CTGCTTTCCA AACATATGAA TGAACAAGC CACACACAGG 1680
GCTCCTTAAG GATAACTAAC ATTTCTATCG ATGACAGTGG GAAGCAGATC TCTTGTGTGG 1740
CGGAAATCT TGTAGAGAAA GATCAAGATT CTGTCAACCT CACTGTGCAT TTTGACCAA 1800
50
CTATCAGATT TCTCGAATCT CCAACCTCAG ACCACCACTG GTGCATTCCA TTCAGTGA 1860
AAGGCAACCC CAACCCAGCG CTTCAGTGGT TCTATAACCG GGCAATATTG AATGAGTCCA 1920
AATACATCTG TACTAAATAA CATGTTACCA ATCACACGGA GTACCAGGCG TGCCTCCAGC 1980
TGGATAATCC CATCACTAG AACAATGGGG ACTACACTCT AATAGCCAAAG AATGAGTATG 2040
GGAAGGATGA GAACAGATT TCTGCTCACT TCAATGGGCTG GCGTGAATTT GACGATGGTG 2100
55
CAACCCCAA TTAATCTGAT GTAATTATG AAGATTATGG AACTGCAGCG AATGACATCG 2160
GGGACACCAC GAACAGAAGT AATGAAATCC CTTCCACAGA CGTCACTGAT AAAACCGGTC 2220
GGGAACATCT CTGCGTCTAT GCTGTGGTGG TGATTGCGTC TGTGGTGGGA TTTGCTTTT 2280
TGGTAATGCT GTTCTGCTAT AAGTTGGCAA GACACTCCAA GTTTGGCATG AAAGGTTTTG 2340
60
TTTTGTTCCT TAAGATCCCA CTGGATGGGT AGCTGAAATA AAGGAAAGA CAGGAAAGG 2400
GGCTGTGGTG CTTGTGGTGT GATGCTGCCA TGTAAAGCTG ACTCCTGGGA CTGCTGTGG 2460
CTTATCCCGG GAAGTGCTGC TTATCTGGGG TTTTCTGGTA GATGTGGGCG GTGTTGGAG 2520
GCTGTACTAT ATGAAGCCTG CATATACTGT GAGCTGTGAT TGGGGAACAC CAATGCAGAG 2580
GTAACCTCTA GGCAGCTAAG CAGCACCTCA AGAAAAATG TTAATTAAT GCTTCTCTTC 2640
TTACAGTAGT TCAATACAA AACTGAAATG AAATCCCAT TGGATTGACT TCTCTCTGA 2700
65
AAAGTGTGCT TTTTGACCTT ACTGGACATT TATTGACTTA ATTGCTCTG TTTATTAATA 2760
TTGACCTGCA AAGTTAAAA AAAATTAAAG TTGAGAACAG GTATAAGTGC ACACCTGAATA 2820
GTCTAATCTA CATGTAACAC ATATTTTATG GTGATTTTCT ATACTCTAAT CAGCACTGAA 2880
TTGAGAGGGT TTGACTTTTT CATCTATAAC ACAGTGACTA AAAGAGTTAA GGGTATATAT 2940
ACCATCACTT TGGGACTTGG TAGTATTATT AAAAGGTTAT TTCCTTCACT GTCAATAAAA 3000
70
GTCCAAATGT TTAGCTTAGG TCTGAGAGTC AAACAATGTT AAGGATTGTC TTAAGTTTCC 3060
TTAGCCAGCA AAACAAAACA AAACAAAACA AACAAATGAA AAACGTTTAA AAAGAAGAAG 3120
AAGAAAAAAG ACAGAACAA GCAGCAACAG CTGTTTGTGT GGGGCTATAG ATTTAAGTTA 3180
GGCATACTCA ATTTTCAAGT AACTAAGAGT GGAATATATG CATATGGTGA AATTATRACC 3240
TTGCCCTTTT TTAATTTGCC TCTGCGATCC ACCTGTCTTT TAGAAGTCTG CCGAGTGAGA 3300
75
AGGCCACAGT ATCTCATGCT GTTTGCATTA CAGAACTGCA GCTTTTCTAC TCTGAAAAGG 3360
CCTGGGAGCA GAATGGCTGG CCTGCTGTGA GCAGGAGAGG AGATTCTAAG AAGGATAGTC 3420
CCCCCTACAA CATACTGTCA TACTGCTGGG TTTTCTATGG TAGGAAAGCT TGTCTGACC 3480
CCAGCAGCAA AGAGGTGGCA GGTGCTAAT GAATATATGC TTTATATGT CTTCTTCTAT 3540
TGCTGAGAGG GCAGCCTTAG AGCTGTGGAT TTCTGCTATC CCGCTGAGTC TGACCCATGG 3600
80
ACACCTGTGT CATTCACCTT AGCATCACAG TGACCTTTGT ATGCTCTGT CAGTCTGTGT 3660
CAGGAGTATG GCTTGTCTTG AAGAGAGGTT TGGCTATCCC CACCCACCC CACCCACCC 3720
TGTTCTTTT TTAATCAGGAG GACTTCAGAG CCAGGCGCTGC AGCATTTTGT TTGAAAACAC 3780
AATCAGCTCT GACAGTTAGA CATGCACACA GACGCCATAG CTGGATTGGA AACATTGATG 3840
TTTAAATAAT TTAATTTTTT TGGAAATAGT TGCACAAATG CTGCAATTTA GCTTTAAGGT 3900
85
TCTATAGATT TTAATAGTAG CCAACACAGT CAGAAACATT GTTTTGAATC CTCTGTAAAC 3960
CAAGGCATTA ATCTTAATAA ACCAGGATCC ATTTAGTATC CACTTGATAT AAAAAGGATA 4020
TCCATAATGA ATATTTTATA CTGCATCCTT TACATTAGCC ACTAAATAGC TTATTGCTTG 4080
ATGAAGACCT TTCACGAAT CCTATGGATT GCAGCATTTT ACTTGGCTAC TTCATACCCA 4140

5 TGCCTTAAAG AGGGGCGATT TCTCAAAAGC AGAAACATGC GSCCAGTTCT CAAGTTTTC 4200
 TCCTAATCC ATTTGAATGT AAGGGCAGCT GGGCCCCAAT GTGGGGAGGT CCGAACATT 4260
 TCTGAATCC CATTTCTTGT TCCGGGCTA AATGACAGTT TCTGTCATTA CTTAGATTCC 4320
 10 GATCTTCC CC AAGGTGTGTG ATTTACAAAG AGGCCAGCTA ATAGCAGAAA TCATGACCCT 4380
 GAAAGAGAGA TGAATTCAC GCTGTGAGCC AGGCAGGAGC TCAGTATGCC AAAGTTTCTT 4440
 GAGAATCAGC CATTTGGTAC AAAAAAGATT TTTAAAGCTT TTATGTTATA CCATGGAGCC 4500
 ATAGAAAGGC TATGGATTGT TTAAGAACTA TTTTAAAGTG TTCAGACCC AAAAAGGAAA 4560
 AATAAAAAA AAGGAATATT TGTACCAAC AGCTAGAAAG ATTGCAAGGT AGATTTTTGT 4620
 15 TTTAAATGG AGAGAAGTGG ACAGATAAGG CCATTTAATA TATCAAAGAT CAGTTGACAT 4680
 CTCCTAGGGA ATGATGAAAA CAGCAGGCTA T

Seq ID NO: 18 Protein sequence:
Protein Accession #: CAAS3571

15 1 11 21 31 41 51
 MSSWIRNHGP AMARLWGFCE LVVGFWRAAF ACPTSCCKCSA SRIWCSDFSP GIVAFPRLEP 60
 NSVDPENITE IFIANQKRL E IINEDDVEAY VGLRNLITVD SGLKFPVAHKA FLKNSNLQHI 120
 NPTRNKLTSL SRKHFRHLDL SELILVGNRP TCSCDIMWK TLQEAQSSPD TODLYCLNES 180
 20 SKNIPLANLQ IPNCGLPAN LAAPNLTVEE GKSLTSLCSV AGDPVPMYV DVGNLVSKHM 240
 NETSHTQGLL RITNISSDDG GKQISCVAEN LVGEDQDSVN LTVHFAPTT FLESPTSDHH 300
 WCIPFTVKNG PKPALQWFYN GAILNESKYI CTKIHVTNHT EYHGLQLDN PTHMNGDYT 360
 LIAKNEYGKD EKQISAHFMG WPGIDDGANP NYPDVIYEDY GTAANDIGDT TNRSNEIPST 420
 25 DVTDKTGREH LSVYAVVVIA SVVGFCLLVM LFLKLKLARHS KFGMKGFVLF HKIPLDG

Seq ID NO: 19 DNA sequence
Nucleic Acid Accession #: NM_000228
Coding sequence: 82-3600

30 1 11 21 31 41 51
 GCTTTGAGC GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAAG TCCTTTCTGG 60
 GGATCACCCC ATTTGGCTGAA GATGAGACCA TTCTTCTCT TGTGTTTTC CCTGCTTGGC 120
 35 CTCTGTCATG CCAACAAGC CTGCTCCCGT GGGGCTCTGT ATCCACCTGT TGGGGACCTG 180
 CTTGTTGGGA GGAACCGGTT TCTCGAGCT TCATCTACCT GTGGACTGAC CAAGCCTGAG 240
 ACCTACTGCA CCGAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCACGGCAG 300
 CCTCACAACT ACTACAGTCA CCGAGTAGAG AATGTGGCTT CATCTCCGG CCCCATGCGC 360
 TGGTGGCAGT CCGAGATGA TGTGAACCTT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420
 40 TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGGCCGCGCG CATGCTGATT 480
 GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCOCAGCTG 540
 ACCTCCACCT TCCTCCGGGT CCGCCAGGGT CGGCCTCAGA GCTGGCAGGA TGTTCGGTGC 600
 CAGTCCCTGC CTCAGAGGCC TAATGCAAGC CTAAATGGGG GGAAGGTCCA ACTTAACCTT 660
 ATGGAATTAG TGTCTGGAT TCCAGCAACT CAAAGTCAA AAATTCAGA GGTGGGGAG 720
 45 ATCACAAACT TGAGAGTCAA TTTCAACAGG CTGGCCCTGT TGGCCCAAG GGGCTACCAC 780
 CCTCCAGCTC CTACTATGCG TGTGTCCAG CTCCGTCTGC AGGGGAGCTG CTTCTGTAC 840
 GGCCATGCTG ATCGCTGGCG ACCCAAGCCT GGGGCTCTGT CAGGCCCTCT CACCGCTGTG 900
 CAGGTCCAGC ATGTCTGTGT CTGCCAGCAC AACACTGCC GCCCAAATTG TGAGCGCTGT 960
 GCACCTTCT ACAACAACCG GCCCTGGAGA CCGCGGAGG GCCAGGACGC CCATGAATGC 1020
 50 CAAAGGTGCG ACTGCAATGG GCACTCAGAG ACATGTCACT TTGACCCGCG TGTGTTTGGC 1080
 GCCAGCCAGG GGCATATGG AGGTGTGTGT GACAAATGCC GGGACCAAC CGAAGGCAAG 1140
 AACTGTGAGC GGTGTGAGT GCACTATTTC CGGAACCGCG GCCCGGGAGC TTCCATTGAG 1200
 GAGACCTGCA TCTCTGCCA GTGTGATCCG GATGGGGCAG TGGCAGGGGC TCCCTGTGAC 1260
 CCAAGTACCG GGCAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320
 55 AAGCGGGCTC TCACTGGACT CACCTACGCC AACCCGCGAG GCTGCCACCG CTGTGACTGC 1380
 AACATCTGG GGTCCCGGAG GACATGCCCG TGTGACGAG AGAGTGGGCG CTGCTTTGT 1440
 CTGCCCAAGC TGTGTGGTCC CAAATGTGAC CAGTGTGCTC CTACCACTG GAAGCTGGCC 1500
 AGTGGCCAGG GCTGTGAACC GTGTGCTGCG GACCCGCACT ACTCCCTCA GCCACAGTG 1560
 CAACCAATTC ACAGGGCAGT GCCCTGTGCG GAAGGCTTTG GTGGCCTGAT GTGCAGCGCT 1620
 60 CGAGCCATCC GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAGG ATGCCAGGCC 1680
 TGTGACTGTG ATTTCCGGGG AACAGAGGGC CCGGGCTGCG ACAAGGCATC AGGCCGCTGC 1740
 CTCTGCCGCG CTGGCTTGAC CGGGCCCCGC TGTGACCACT GCCAGCGAGG CTACTGCAAT 1800
 CGCTACCCCG TGTGCTGTGG CTGCCACCTT TGCTTCCAGA CCTATGATGC GGACCTCCGG 1860
 GAGCAGGCCG TGGCTTTTGG TAGACTCCGC AATGCCACCG CCAGCCTGTG GTGAGGGCCT 1920
 65 GGGCTGGAGG ACCGTGGCCT GGCCTCCCGG ATCTAGATG CAAAGAGTAA GATTGAGCAG 1980
 ATCCAGAGCA TTCTCAGCAG CCCCGCAGTC ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT 2040
 GCCATCCTCT CCCTCAGGCG AACTCTCAG GGCCTGCAGC TGGATCTGCC CCTGGAGGAG 2100
 GAGAGCTTGT CCCTTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCCTT 2160
 ACTATGTATC AGAGGAAGAG GGAGCAGTTT GAAAAAATAA GCAGTGTCTG TCCTTCAGGA 2220
 70 GCCTTCOGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGTCTA GCAGGTCTCC 2280
 GACAGCTGCG GCTTTTGGGA CCAGCTCAGG GACAGCCGGA GAGAGGCAGA GAGGCTGGTG 2340
 CGGCAGGGCG GAGGAGGAGG AGGCACCGGC AGCCCCAAGC TTGTGGCCCT GAGGCTGGAG 2400
 ATGTCTTCGT TGCTGACCT GACACCCACC TTCAACAAGC TCTGTGGCAA CTCAGGCGAG 2460
 ATGGCTTGCA CCCCATATC ATGCCCTGGT GAGCTATGTC CCCAAGACAA TGGCACAGCC 2520
 75 TGTGGCTCCC GCTGCAAGGG TGTCTTCCC AGGGCCGGTG GGGCCTTCTT GATGGCGGGG 2580
 CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCCAGCTCC AGCGGACCA GAGATGATT 2640
 AGGGCAGCCG AGGAATCTGC CTACAGATT CAATCCAGTG CCCAGCGCTT GGAGACCCAG 2700
 GTGAGCGCCA CGCGCTCCCA GATGGAGGAA GATGTACAGC GCACACGGCT CCTAATCCAG 2760
 CAGGTCCGGG ACTTCCCTAAC AGACCCCGAC ACTGATGACG CCATATCCA GGAGGTGAGC 2820
 80 GAGGCGGTGC TGGCCCTGTG GCTGCCCACA GACTCAGCTA CTGTTCTGCA GAAGATGAAT 2880
 GAGATCCAGC CCACTTGCAG CAGGCTCCCC AACGTGGACT TGGTGTCTGT CCAGACCAAG 2940
 CAGSACATTG CGCTGCCCGG CCGGTTGCAG GCTGAGGCTG AGGAAGCCAG GAGCCGAGCC 3000
 CATGCACTGG AGGGCCAGGT GGAAGATGTG GTTGGGAACC TGGCGCAGGG GACAGTGGCA 3060
 CTCAGGAGAG CTCAGAGAC CATGCAAGG ACCAGCCGCT CCTTCGGCTT TATCCAGGAC 3120
 85 AGGGTTGCTG AGGTTCAGCA GGTACTGCGG CCAGCAGAAA AGCTGGTGAC AAGCATGACC 3180
 AAGCAGCTCG GTGACTCTG GACACGGATG GAGGAGCTCC GCCCAAGC CCGGCAGCAG 3240
 GGGCAGAGG CAGTCCAGGC CCAGCAGCTT GCGGAAGGTG CCAGCGAGCA GGCATTGAGT 3300
 GCCCAAGAGG GATTTGAGAG AATAAACCAA AAGTATGCTG AGTGAAGGA CCGGTTGGGT 3360

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

CAGAGTTCCA TGCTGGGTGA GCAGGGTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG 3420
 GAGCTGTTTG GGGAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480
 CTGGGGGGCA GCCAGGCCAT CATGCTGC GC GCGCGGACC TGACAGGACT GGAGAAGCGT 3540
 GTGGAGCAGA TCCGTGACCA CATCAATGGG CGCGTGCTCT ACTATGCCAC CTGCAAGTGA 3600
 TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCCGCC TTTGCTTTTG GTTGGGGCCA 3660
 GATTGGGTTG GAATGCTTTC CATCTCCAGG AGACTTTCAT GCAGCCTAAA GTACAGCCTG 3720
 GACCAACCCCT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAAT 3780
 GGGACAGTTA CACTTGACAG ACAAGATGG TGGAGATTGG CATGCCATTG AACTTAAGAG 3840
 CTCTCAAGTC AAGGAAGCTG GGTGGGCGAG TATCCCCGCG CTTTAGTTCT CCACTGGGGA 3900
 GGAATCCTGG ACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAAATGAA AAGCCAAATA 3960
 AAAATCTTTG G

Seq ID NO: 20 Protein sequence:
 Protein Accession #: NP_000219

1 11 21 31 41 51
 MRFFLLCFA LPGLLHAQQA CSRGACYPPV GDLLVGRTRF LRASSTCGLT KPETYCTQYG 60
 EQMKCKCKD SRQPHNYSH RVENVASSG PMRNWQSQND VNPVSLQLDL DRRPQLQEVN 120
 MEFGQPMFAG MLIERSSDFG KTRVYQYLA ADCTSTFPRV RQGRFPQSWQD VRCQSLPQRP 180
 NARLNGKGVQ LNLMDLVSGI PATQSQKIQE VGEITNLRVN PTRLAPVPQR GYHPFSAYYA 240
 VSQRLRQASC FCHGHADRCF PKPGASAGPS TAVQVHDVCV CQHNTAGPNC ERCAFFYNRR 300
 PWRPAEGQDA HEQRCDCMNG HSETCHFDPA VFAASQGAYG GVCNCRDHT ESKNCERCQL 360
 HYFRNRPPGA SIQETCISCE CDPDGAVPGA PCDPVTGQCV CKEHVQGERC DLCKPGFTGL 420
 TYANPQGERH CDNIIIGSRR DMPCEESGR CLCLPNVVGK KCDQCAPYHW KLASGQGCPE 480
 CACDPHNSPQ PTVPQVRAV PCREGGFGLM CSAARIRQCP DRTYGDVATG CRACDCDFRG 540
 TEGPGCDKAS GRCLCRPGLT GPRCDQCRG YCNRYFVCVA CHPCFQTYDA DLREQALRFG 600
 RLERNATSLW SGPGLEDRLG ASRILDAKSK IEQIRAVLSS PAVTEQEVQA VASAILSLRR 660
 TLQGLQLDLP LEEETLSLPR DLESLDSPFN GLLTMYQRKR EQFEKISSAD PSGAFRMLST 720
 AYEQSAQAAQ QVSDSSRLLD QLRDSRREAE RLVRQAGGGG GTGSPKLVAL RLEMSSLPDL 780
 TPTFNKLCGN SRQMACTPIS CPHELCPQDN GTACGSRRCRG VLPRAAGAPL MAGQVABQLR 840
 GFNAQLQRTM QTRAAEESA SIQSSAQLR ETQVSASRSQ MEEDVRRTRL LIQQVRDFLT 900
 DPDTDAATIQ EVSEAVLALW LPTDSATVLQ KMNEIQAIAA RLPNVDLVLS QTKQDIARAR 960
 RLQAEAEAR SRAHVEGVQ EDVVGNLRRQ TVALQEAQDT MQGTSRSLRL IQDRVAEVQQ 1020
 VLRPAEKLVY SMTKQLGDFM TRMEELRHQA RQQGAFAVQA QQLAEGSAQ ALSAQEGFPR 1080
 IKQKYAELEK RLQSSMLGE QGARIQSVKT EAEELPGETM EMDRMKDME LELLRGSQAI 1140
 MLRSADLTGL EKRVEQIRDH INGRVLYYAT CK

Seq ID NO: 21 DNA sequence
 Nucleic Acid Accession #: NM_003722
 Coding sequence: 145-1491

1 11 21 31 41 51
 TCGTTGATAT CAAAGACAGT TGAAGGAAAT GAATTTTGAA ACTTCACGGT GTGCCACCCT 60
 ACAGTACTGC CCTGACCCTT ACATCCAGCG TTTCGTAGAA ACCCAGCTCA TTTCTCTTGG 120
 AAAGAAGAGT ATTACCGATC CACCATGTCC CAGAGCACAC AGACAAATGA ATTCTCTCAGT 180
 CCAGAGGTTT TCCAGCATAT CTGGGATTTT CTGGAACAGC CTATATGTTT AGTTTACGCC 240
 ATTGACTTGA ACTTTGTGGA TGAACCATCA GAAGATGGTG CGACAAACAA GATTGAGATT 300
 AGCATGGACT GTATCCGCAT GCAGGACTCG GACCTGAGTG ACCCCATGTG GCCACAGTAC 360
 ACGAACCTGG GGCCTCTGAA CAGCATGGAC CAGCAGATTC AGAACGGCTC CTGCTCCACC 420
 AGTCCCTATA ACACAGACCA CGCGCAGAAC AGCGTCACGG CGCCCTCGCC CTACGCACAG 480
 CCGAGCTCCA CCTTCGATGC TCTCTCTCCA TCACCCGCCA TCCCTCCAA CACCGACTAC 540
 CCGAGCCCGC ACAGTTTCCA CGTGTCTTTC CAGCAGTCCA GCACCGCCAA GTGCGCCACC 600
 TGGAGCTATT TTGGCACTGA GAAGAAACTC TACTGCCAAA TTGCAAAAGC ATGCCCCATC 660
 CAGATCAAGG TGATGACCCC ACCTCCTCAG GGAGCTGTGA TCGCGCCCAT GCCTGTCTAC 720
 AAAAAAGCTG AGCAGCTCAC GGAGGTGGTG AAGCGGTGCC CCAACCATGA GCTGAGCCGT 780
 GAATTCACAG AGGACAGAT TCCCTCTCCT AGTCAITTTA TTGAGTAGA GGGGAACAGC 840
 CATGCCAGT ATGTAGAAGA TCCCATCACA GGAAGACAGA GTGTGCTGGT ACCTTATGAG 900
 CCACCCCGAG TTGGCACTGA ATTCACGACA GTCTTGTA TAATTCATGT TAACAGCAGT 960
 TGTGTGGAG GGTGAACCG CCGTCCAAAT TTAATCATTT TTAATCTGGA AACCAGAGAT 1020
 GGGCAAGTCC TGGGCCGAGC CTGCTTTGAG GCCCGGATCT GTGCTTGCCC AGGAAGAGAC 1080
 AGGAAGGCGG ATGAAGATAG CATCAGAAAG CAGCAAGTTT CGACAGTAC AAAGAAGCGT 1140
 GATGGTACGA AGCGCCCGTT TGTGAGAAC ACACATGGTA TCCAGATGAC ATCCATCAAG 1200
 AAACGAAGAT CCCAGATGA TGAACGTGTA TACTTACCAG TGAGGGGCGG TGAGACTTAT 1260
 GAAATGCTGT TGAAGATCAA AGAGTCCCTG GAATCTATGC AGTACCTTCC TCAGCACACA 1320
 ATTGAAACGT ACAGGCAACA GCAACAGCAG CAGCACCAGC ACTTACTTCA GAAACATCTC 1380
 CTTTCAGCCT GCTTCAGGAA TGAGCTTGTG GAGCCCCGGA GAGAAACTCC AAAACATCTC 1440
 GAGCTCTTCT TTAGACATTC CAAGCCCCCA AACCGATCAG TGTACCCATA GAGCCCTATC 1500
 TCTATATTTT AAGTGTGTGT GTTGTATTTC CATGTGTATA TGTGAGTGTG TGTGTGTGTA 1560
 TGTGTGTGCG TGTGTATCTA GCCCTCATAA ACAGGACTTG AAGACACTTT GGCTCAGAGA 1620
 CCCAATCTGT CAAAGGCACA AAGCCACTAG TGAGAGAAAT TTTTGAAGGG ACTCAAACTC 1680
 TTACAAGAAA GGATGTTTTC TGCAATTTT GTATCCTTAG ACCCGCCATT GGTGGGTGAG 1740
 GAACCACTGT GTTGTCTGT GAGCTTCTG TGTCTTCTG GGAGGGAGGG GTCAAGTGGG 1800
 GAAAGGGGCA TTAGATGTTT TATTGGAACC CTTTCTGTC TCTCTCTGTT GTTTTCTTAA 1860
 AATTCACAGG GAAGCTTTTG AGCAGGTCTC AAATCTAAGA TGTCTTTTGA AGAAAAGGAG 1920
 AAAAAAGTTG TTATTGTCTG TGCATAAGTA AGTTGTAGGT GACTGAGAGA CTCAGTCAGA 1980
 CCCCTTTAAT GCTGGTCATG TAATAATATT GCAAGTAGTA AGAAACGAAG GTGTCAAGTG 2040
 TACTGCTGGG CAGCGAGGTG ATCATTACCA AAAGTAATCA ACTTGTGGG TGGAGAGTTC 2100
 TTTGTGAGAA CTGTCAATT TGTGTCTCTC CCTCATGTG TAGGTAGAAC ATTTCTTAAT 2160
 GCTGTGTACC TGCCCTCTGC ACTGTATGTT GGCATCTGTT ATGCTAAAGT TTTTCTTGTA 2220
 CATGAACCTC TGAAGACCT ACTACAAAA AACTGTGTGT TGGCCCCCAT AGCAGGTGAA 2280
 CTCAATTTGT GCTTTTAATA GAAAGACAAA TCCACCCAG TAATATTGCC CTTACGTAGT 2340
 TGTTTACCAT TATTCAAGC TCAAAATAGA ATTTGAAGCC CTCTCACAAA ATCTGTGATT 2400
 AATTTGCTTA ATTAGAGCTT CTATCCCTCA AGCCTACCTA CCATAAAACC AGCCATATTA 2460
 CTGATACTGT TCAAGTCATT TAGCCAGGAG ACTTACGTTT TGAGTAAGTG AGATCCAAGC 2520
 AGACGTGTTA AATCAGCAC TCCTGGACTG GAAATTAAG ATTGAAGGG TAGACTACTT 2580

TTTCTTTT TACTCAAAAG TTTAGAGAAT CTCTGTTTCT TTCCATTTTA AAAACATATT 2640
 TTAAGATAAT AGCATAAAGA CTTTAAAAAT GTTCTCTCCC TCCATCTTCC CACCCACAGT 2700
 CACCAGCACT GTATTTTCTG TCACCAAGAC AATGATTTCT TGTATTAGAG GCTGTTGCTT 2760
 TTGTGGATGT GTGATTTTAA TTTTCAATAA ACTTTTGCAT CTGTGTTTAA AAGAAA

Seq ID NO: 22 Protein sequence:
 Protein Accession #: NP_003713

1 11 21 31 41 51
 MSQSTQTNEF LSPEVFQHIW DFLEQPICSV QPIDLNFVDE PSEDGATNKI EISMDCIRMQ 60
 DSDLSDPMWP QYTNLGLLNS MDQQIQNGSS STSPYNTDHA QNSVTAPSPY AQPSSSTFDAL 120
 SPSPAIPTNT DYPGPHSFDV SFQSSSTAKS ATWTYSTELK KLYCOIAKTC PIQIKVMTFP 180
 PQGAVIRAMP VYKQAEHVTE VVKRCPNHEL SREFNEGQIA PPSHLIRVEG NSHAQYVEDP 240
 ITGRQSVLVP YEPPQVGTGF TTVLYNFMCN SSCVGGMNRR PILIIVTLET RDGQVLGRRR 300
 FEARICACPG RDRKADSDSI RKQQVSDSTK NGDGTKRPPR QNTHGIQMTS IKRRRSPDDE 360
 LLYLPVRGRE TYEMLLKIKE SLELMQVLPQ HTIETRYQQQ QQQHQLLQK HLLSACPRNE 420
 LVEPRRETPK QSDVFRHSHK PPNRSVYP

Seq ID NO: 23 DNA sequence
 Nucleic Acid Accession #: NM_001944.1
 Coding sequence: 84-3083

1 11 21 31 41 51
 TTTTCTTAGA CATTAACTGC AGACGGCTGG CAGGATAGAA GCACGGGCTC ACTTGGACTT 60
 TTTACACAGG GAAATCAGAG ACAATGATGG GGCTCTTCCC CAGAACTACA GGGGCTCTGG 120
 CCATCTTCGT GGTGGTCATA TTGGTTCATG GAGAATTGGG AATAGAGACT AAAGGTCAAT 180
 ATGATGAAGA AGAGATGACT ATGCAACAAG CTAAAAGAAG GCAAAAACGT GAATGGGTGA 240
 AATTGTCCAA ACCCTGCAGA GAAGGAGAAG ATAACCTCAA AAGAAACCCA ATTGCCAAGA 300
 TTACTTCAGA TTACCAAGCA ACCCAGAAAA TCACCTACOG AATCTCTGGA GTGGGAATCG 360
 ATCAGCCGCC TTTTGGGAATC TTTGTTGTTG ACAAACACAC TGGAGATATT AACATAACAG 420
 CTATAGTCGA CGGGGAGGAA ACTCCAAGCT TCCTGATCAC ATGTGGGGCT CTAATATGCC 480
 AAGGACTAGA TGTAGAGAAA CCACCTTATC TAACGGTTAA AATTTTGGAT ATTAATGATA 540
 ATCTCCAGT ATTTTCACAA CAAATTTTCA TGGGTGAAAT TGAAGAAAT AGTGCCTCAA 600
 ACTCACTGGT GATGATACTA AATGCCACAG ATGCAGATGA ACCAAACCCAC TTGAATTCTA 660
 AAATTGCCCT CAAATTTGTC TCTCAGGAAC CAGCAGGCAC ACCCATGTTC CTCCTAAGCA 720
 GAAACACTGG GGAAGTCCGT ACTTTGACCA ATTTCTCTGA COGAGAGCAA GCTAGCAGCT 780
 ATCGTCTGAT TGTGAGTGGT GCAGACAAAG ATGGAGAAGG ACTATCAACT CAATGTGAAT 840
 GTAATATTAA AGTGAAGAT GTCAACGATA ACTTCCCAAT GTTTAGAGAC TCTCAGTATT 900
 CAGCAGTAT TGAAGAAAT ATTTTAAAGT CTGAATTACT TCGATTTCAA GTACAGATT 960
 TGGATGAAGA GTACACAGAT AATTGGCTTG CAGTATATTT CTTTACCTCT GGGAAATGAAG 1020
 GAAATGTGTT TGAATACAA ACTGATCCTA GAACTAATGA AGGCATCCTG AAAGTGGTGA 1080
 AGGCTCTAGA TGTAGAACAA CTACAAAGCG TGAACCTTAG TATTGCTGTC AAAAACAAAG 1140
 CTGAATTTCA CCAATCAGTT ATCTCTGAT ACCGAGTTCA GTCAACCCCA GTCACAATTC 1200
 AGGTAATATA TGAAGAGAA GGAATTGCAT TCGTCTGTC TTCCAGACA TTTACTGTGC 1260
 AAAAAGGCAT AAGTAGCAAA AAATTGGTGG ATTATATCCT GGGAACATAT CAGGCCATCG 1320
 ATGAGGACAC TAACAAGCT GCCTCAAATG TCAATATGTT CATGGGACGT AACGATGGTG 1380
 GATACCTAAT GATTGATTCA AAAACTGCTG AAATCAAATT TGTCAAAAT ATGAACCGAG 1440
 ATTCTACTTT CATAGTTAAT AAAACAATCA CAGCTGAGGT TCTGGCCATA GATGAATACA 1500
 CGGGTAAAAA TTCTACAGGC ACGGTATATG TTAGAGTACC CGAATTTCAAT GACAATTGTC 1560
 CAACAGCTGT CCTCGAAAAA GATGCAGTTT GCAGTCTTTC ACCTTCCGTG GTTGTCTCCG 1620
 CTAGAACACT GAATAATAGA TACACTGGCC CCTATACATT TGCACTGGAA GATCAACCTG 1680
 TAAAGTTGCC TCTGCTGTTG AGTATCACAA CCTTCAATGC TACCTCGGCC CTCTCTCAGG 1740
 OCCAGGAACA GATACCTCCT GGAGTATACC ACATCTCCTT GGTACTTACA GACAGTCAGA 1800
 ACAATCGGTG TGAGATGCCA CGCAGCTTGA CACTGGAAGT CTGTCACTGT GACAACAGGG 1860
 GCATCTGTGG AACTTCTTAC CCAACCAAA GCCCTGGGAC CAGGTATGGC AGGCGGCACT 1920
 CAGGAGGCT GGGGCTGCCC GCCATCGGCC TGCTGCTCCT TGGTCTCCTG CTGCTGCTGT 1980
 TGGCCCCCTT TCTGCTGTTG ACCTGTGACT GTGGGGCAGG TTCTACTGGG GAGGTGACAG 2040
 GTGGTTTTAT CCACTGTCTT GATGGCTCAG AAGGAACAAT TCATCAGTGG GGAATTGAAG 2100
 GAGCCCATCC TGAAGACAAG GAAATCACAA ATATTTGTGT GCCTCCTGTA ACAGCCAATG 2160
 GAGCGGATTT CATGGAAGTT TCTGAAGTTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220
 TGGAAAGCAC TTCAGGAATG GAAATGACCA CTAAGCTTGG AGCAGCCACT GAATCTGGAG 2280
 GTGCTGCAGG CTTTGCACAA GGGACAGTGT CAGGAGCTGC TTCAGGATTC GGAGCAGCCA 2340
 CTGAGTTGG CATCTGTTC TCAGGGCAGT CTGGAACCAT GAGAACRAGG CATTCCACTG 2400
 GAGGAACCAA TAAGGACTAC GCTGATGGGG CGATAAGCAT GAATTTCTG GACTCCTACT 2460
 TTTCTCAGAA AGCATTGTCC TGTGCGGAGG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520
 TGTGTATCTA TGATAATGAA GGCAGCATG CCACCTGGTT TCTGTGGGC TCGGTGGGTT 2580
 GTTGCAGTTT TATTGCTGAT GACCTGGATG ACAGCTTCTT GGACTCACTT GGACCCAAAT 2640
 TTAATAAACT TGCAGAGATA AGCCTTGGTG TTGATGGTGA AGGCAARGAA GTTCAGCCAC 2700
 CCTCTAAAGA CAGCGGTTAT GGGATTGAAT CCTGTGGCCA TOCCATAGAA GTCCAGCAGA 2760
 CAGGATTTGT TAAGTCCCAG ACTTTGTGAG GAAGTCAAGG AGCTTCTGCT TTGTCCGCTT 2820
 CTGGGTTCTG CCAGTCAGCT GTTTCATATC CTGACCCCTCT GCAGCATGGT AACTATTTAG 2880
 TAAAGGAGAC TTAATCGGCT TCTGGTTCCC TCGTGCRAAC TTCCACTGCA GGCCTTGATC 2940
 CACTTCTCAC ACAAATGTG ATAGTGACAG AAAGGGTGAT CTGTCCCAT TCCAGTGTTC 3000
 CTGGCAACCT AGCTGGCCCA ACGCAGCTAC GAGGGTCACA TACTATGCTC TGTACAGAGG 3060
 ATCCTTGCTC CGCTCTAATA TGACCAGAAAT GAGCTGGAAT ACCACACTGA CCAATCTGG 3120
 ATCTTTGGAC TAAAGTATTC AAAATAGCAT AGCAAAAGCTC ACTGTATTGG GCTAATAATT 3180
 TGGCACTTAT TAGCTTCTCT CATAAACTGA TCACGATTAT AAATTAATG TTTGGGTTCA 3240
 TACCCCAAAA GCAATATGTT GTCACTCCTA ATTCTCAAGT ACTATTCAAA TTGTAGTAAA 3300
 TCTTAAAGTT TTTCAAAACC CTAAATCAT ATTGCG

Seq ID NO: 24 Protein sequence:
 Protein Accession #: NP_001935.1

1 11 21 31 41 51

	MMGLPFRRTTG	ALAIFVNVIL	VHGLRIETK	QGYDEEEMTM	QQAQRKQRE	VVKFAKPCRE	60
	GEDNSIKRNP	AKITSDYQAT	OKITYRISGV	GIDQPPFGIF	VVDKNTGDIN	ITAIVDREET	120
	PSFLITCRAL	NAQGLDVEKP	LILTVKILDI	NDNPPVFSQQ	IFMGEIEENS	ASNSLVMILN	180
5	ATDADEPNHL	NSKIAFKIVS	QEPAGTMMFL	LSRNTGEVRT	LTNSLDREQA	SSYRLVVSQA	240
	DKDGEGLSTQ	CECNIKVKDV	NDNPFMRDS	QYSARIEENI	LSSELLRFQV	TDLDEEYTDN	300
	WLAVYFPTGG	NEGNWPEIQ	DPRNNEGILK	VVKALDYEQL	QSVKLSIAVK	NKAEPHQSVI	360
	SRVRVQSTPV	TIQVINVRG	IAPRPASKTF	TVQKGISSEK	LVDYILGTQY	AIDEDTNKAA	420
	SNVKYVMGRN	DGGLVIMDSK	TAEIKFVKNM	NRDSTFIVNK	TITAEVLAD	EYTGKTSTGT	480
10	VYVRVPDFND	NCPTAVLEKD	AVCSSSPSVV	VSARTLNNRY	TGPTYFALED	QPVKLPAVWS	540
	ITTLNATSL	LRAQEQIPPG	VYHISLVLT	SQNNRCMPR	SLTLEVCQCD	NRGICGTSYP	600
	TTSPGTRYGR	PHSGRLGPA	IGLLLLLGLL	LLAPLPLLLT	CDCGAGSTGG	VTGGFIPVPD	660
	GSEGTIHQWG	IEGAPHEDE	ITNICVPPVT	ANGADFMESS	EVCTNTYARG	TAVEGTSQME	720
	MTTKLGAATE	SGGAAGFATG	TVSGAASGFG	AATGVGICSS	QSGTMRTRH	STGGTNKDYA	780
15	DGALSMNFD	SYFSQKAFAC	AEEDDGQEAN	DCLLIYDNEG	ADATGSPVGS	VGCCSFIADD	840
	LDDSFILDSG	PKFKLAEIS	LGVDGEGKEV	QPPSKDSGYG	IESCGHPIEV	QQTGFVKCQT	900
	LSGSQASAL	SASGSVQPAV	SIPDPLQHGN	YLVTEYSAS	GSLVQPSTAG	FDPLLTQNV	960
	VTERVICPIS	SVPGNLAGPT	QLRGSHTMLC	TEDPCSRLI			

Seq ID NO: 25 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 56-1642

	1	11	21	31	41	51	
25	AGTATCCCAG	GAGGAGCAAG	TGGCACGTCT	TCGGACCTAG	GCTGCCCCCTG	CCGTCTATGTC	60
	GCAAGGATC	CTTTCTCCGC	CAGCGGGCTT	GCTGTCCGAT	GACGATGTGG	TAGTTTCTCC	120
	CATGTTTQAG	TCCACAGCTG	CAGATTGCGG	GTCTGTGCTA	CGCAAGAAC	TGCTATCAGA	180
	CTGCTCTGTC	GTCTCTACCT	CCCTAGAGGA	CAAGCAGCAG	GTTCCATCTG	AGGACAGTAT	240
30	GGAGAAGGTC	AAAGTATACT	TGAGGGTTAG	GCCCTTGTTA	CCTTCAGAGT	TGGAACGACA	300
	GGAAAGATCG	GGTGTGTGCC	GTATTGAGAA	TGTGGAGACC	CTTGTCTTAC	AAGCACCCAA	360
	GGACTCTTTT	GCCCTGAAGA	GCAATGAACG	GGGAATTGGC	CAAGCCACAC	ACAGGTTTAC	420
	CTTTTCCCAG	ATCTTTGGGC	CAGAAGTGGG	ACAGGCATCC	TTCCTCAACC	TAACTGTGAA	480
	GGAGATGTA	AAGGATGTAC	TCAAAGGGCA	GAATCTGGCT	ATCTATACAT	ATGGAGTCAC	540
35	TAACTCAGGG	AAAACCCACA	CGATTCAAGG	TACCATCAAG	GATGGAGGGA	TCTCCCCCG	600
	GTCCTTGGCG	CTGATCTTCA	ATAGCCTCCA	AGGCCAACTT	CATCCAACAC	CTGATCTGAA	660
	GCCCTGTCTC	TCCAATGAGG	TAATCTGGCT	AGACAGCAAG	CAGATCCGAC	AGGAGGAAAT	720
	GAAGAAGCTG	TCCCTGCTAA	ATGGAGGCCT	CCAAGAGGAG	GAGCTGTCCA	CTTCTCTGAA	780
	GAGGAGTGTC	TACATCGAAA	GTCCGATAGG	TACCAGCACC	AGCTTCGACA	GTGGCATTGC	840
40	TGGGCTCTCT	TCTATCAGTC	AGTGTACCAG	CAGTAGCCAG	CTGGATGAAA	CAAGTCATCG	900
	ATGGGACACG	CCAGCACTG	CCCACTATCC	TGTCCCGCCA	AACATTCGCT	TCTCCATCTG	960
	GATCTCATTC	TTTGAGATCT	ACAACGAAC	GCTTTATGAC	CTATTAGAAC	CGCCTAGCCA	1020
	ACAGCGCAAG	AGGCAGACTT	TGCGGCTATG	CGAGGATCAA	AATGGCAATC	CCTATGTGAA	1080
	AGATCTCAAC	TGGATTCTAT	TGCAAGATGC	TGGAGAGGCC	TGGAAGCTCC	TAAAGTGGG	1140
45	TCGTAAGAAC	CAGAGCTTTG	CCAGCACCCA	CTCAACCCAG	AACCTCCAGC	CGAGTCACAG	1200
	CATCTTCTCA	ATCAGGATCC	TACACCTTCA	GGGGGAAGGA	GATATAGTCC	CCAAGATCAG	1260
	CGAGCTGTCA	CTCTGTGATC	TGCGTGGCTC	AGAGCGCTGC	AAAGATCAGA	AGAGTGGTGA	1320
	ACGGTTGAAG	GAAGCAGGAA	ACATTAAAC	CTCTCTACAC	ACCCCTGGGC	GCTGTATTGC	1380
	TGCCCTTCCT	CAAAACCCAG	AGAACCGGTC	AAAGCAGAAC	CTGGTTCCCT	TCCGTGACAG	1440
50	CAAGTTGACT	CGAGTGTCTC	AAGGTTTCTT	CACAGGCCGA	GGCGGTTCCCT	GCATGATTGT	1500
	CAATGTGAAT	CCCTGTGCAT	CTACCTATGA	TGAAACTCTT	CATGTGGCCA	AGTCTCAGC	1560
	CATTGCTAGC	CAGGTGACTT	GTGCTATGCC	CACCTATGCA	ACTGGGATTC	CCATCCCTGC	1620
	ACTGTTTCAT	CAAGGAACAT	AGTCTTCAGG	TATCCCCCAG	CTTAGAGAAA	GGGGCTAAGG	1680
	CAGACACAGG	CCCTGTATGAT	GATATTGAAA	ATGAAGCTGA	CATCTCCATG	TATGGCAAGG	1740
55	AGGAGCTCCT	ACAAGTTGTG	GAAGCCATGA	AGACACTGCT	TTTGAAGGAA	CGACAGGAAA	1800
	AGCTACAGCT	GGAGATGCAT	CTCCGAGATG	AAATTTGCAA	TGAGATGGTA	GAACAGATGC	1860
	AACAGCGGGA	ACAGTGGTGC	AGTGAACATT	TGACACCCA	AAAGSAACTA	TTGGAGGAAA	1920
	TGTATGAAGA	AAAACCTAAT	ATCCTCAAGG	AGTCACTGAC	AAGTTTITAC	CAAGAAGAGA	1980
	TTAGAGAGCG	GGATGAAAAG	ATTGAAGAGC	TAGAAGCTCT	CTTGACAGAA	GCCAGACAAC	2040
60	AGTCAGTGCT	CCATCTGGC	TCAGGGTCTG	AATTGGCCCT	ACGGCGGCTC	CAAGGTTTGG	2100
	CAGCTTCTGC	CTCCACCCAG	CAGCTTCAGG	AGGTTAAAGC	TAAATTACAG	CAGTGCAAA	2160
	CAGAGCTAAA	CTCTACCACT	GAAGAGTTGC	ATAAGTATCA	GAAATGTGTA	GAACCACTCC	2220
	CCTCAGCCAA	GCCCTTCACC	ATTGATGTGG	ACAAGAAAGT	AGAAGAGGGC	CAGAAGAATA	2280
	TAAGGCTGTT	GCGGACAGAG	CTTCAGAAAC	TTGGTGAGTC	TCTCCAATCA	CGACAGAGAG	2340
65	CTTGTGTTCA	CAGCACTGGG	GCAGGAAAAC	TTGGTCAAGC	CTTGACCACT	TGTGATGACA	2400
	TCTTAATCAA	ACAGGACCCG	ACTCTGGCTG	AACTGCGAAA	CAACATGGTG	CTAGTGAAC	2460
	TGGACCTTGG	GAAGAAGGCA	GCATGTATTG	CTGAGCAGTA	TCATCTGTG	TTGAACTCTC	2520
	AAGGCCAGGT	TTCTGCCAAA	AAGCGCCTTG	GTACCAACCA	GGAAATACAG	CAACCAAAAC	2580
	AACAAACACC	ASGGAAGAAA	CCATTCTCTC	GAAATTTACT	TCCCCGAAAC	CCAACCTGCC	2640
70	AAAGCTCAAC	AGACTGCGAC	CCTTATGCC	GGATCCTACG	CTCACGGCGT	TCCCCTTTAC	2700
	TCAAATCTGG	GCCTTTTGGC	AAAAGTACT	AAGGCTGTGG	GGAAAGAGAA	GAGCAGTCAT	2760
	GGCCCTGAGG	TGGGTGAGCT	ACTCTCTCTG	AGAAATAGGT	CTCTTTTATG	CTTTACCAT	2820
	TATCAGGAAT	TATATCCAGG	ATGCAATACT	CAGACACTAG	CTTTTCTCTC	ACTTTTGTAT	2880
	TATAACCAAC	TATGTAATCT	CATGTTGTTG	TTTTTTTTTA	TTTACTTATA	TGATTTCTAT	2940
75	GCACACAAA	ACAGTTATAT	TAAAGATATT	ATTGTTTACA	TTTTTTTATG	AATTCCAAAT	3000
	GTAGCAAAAT	CATTAAACAA	AATTATAAAA	GGGACAGAAA	AA		

Seq ID NO: 26 Protein sequence:
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
80	MSQGISPPA	GLLSDDVVV	SPMFESTAAD	LGSVVRKNLL	SDCSVVSTSL	EDKQVPSSE	60
	SMEVKVYLR	VRPLLPSELE	RQEDQGCVR	ENVETVLQQA	PKDSFALKSN	BRGIGQATHR	120
85	FTFSQIFGPE	VQASFPNLT	VKEMVKDVLK	GQNWLIYTYG	VINSKGTETI	QGTIKDGGIL	180
	PRSLALIFNS	LQGLLHPTPD	LKPLLSNEVI	WLDKQIRQRE	EMKLSLLNG	GLQEEELSTS	240
	LKRSVYIBSR	IGTSTSPDSG	IAGLSSISQC	TSSSQLEDS	HRWAQPDTPA	LPVPANIRPS	300

IWISFFEIYN ELLYDLLEPP SQQRKRQTLR LCEQNGNPNY VKDLNWIHVQ DAEAWKLLK 360
 VGRKQSPAS THNQNSSRS HSIFSIIRILH LQEGDIVPK ISELSCDLA GSERCKDQKS 420
 GERLKEAGNI NLSLHTLGRIC IALRQNNQN RSKQNLVPPR DSKLTRVFG FFGTGRGRSCM 480
 IVNVNFCST YDETLHVAKP SAIASQVTC CPTATGIPI PALVHQGT

Seq ID NO: 27 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 13-1424

10 1 11 21 31 41 51
 TAGAAGTTTA CAATGAAGTT TCTTCTAATA CTGCTCCTGC AGGCCACTGC TTCTGGAGCT 60
 CTTCCTCTGA ACAGCTCTAC AAGCCTGGAA AAAAATAATG TGCTATTGGG TGAAAGATAC 120
 TTAGAAAAAT TTTATGGCCT TGAGATAAAC AAACCTCCAG TGACAAAAAT GAAATATAGT 180
 15 GGAAACTTAA TGAAGGAAAA AATCCAAGAA ATGCAGCACT TCTTGGGTCT GAAAGTGACC 240
 GGGCAACTGG ACACATCTAC CCTGGAGATG ATGCACGCAC CTGGATGTGG AGTCCCGAT 300
 GTCCATCTAT TCAGGGAAAT GCCAGGGGGG CCGTATGGA GGAAACATTA TATCACTAC 360
 AGAATCAATA ATTACACACC TGACATGAAC CGTGAGGATG TTGACTACGC AATCCGGAAA 420
 GCTTCCCAAG TATGGAGTAA TGTTACCCCT TTGAAATTCA GCAAGATTAA CACAGGCATG 480
 20 GCTGACATTG TCGTGTGTTT TGCCGTGGA GCTCATGGAG ACTTCCATGC TTTTGTATGGC 540
 AAAGGTGAAA TCCTAGCCCA TGCTTTTGA CCTGGATCTG GCATTGGAGG GGATGCACAT 600
 TTGATGAGG ACAGAACTCT GACTACACAT TCAGGAGGCA CAAACTTGT CCTCACTGCT 660
 GTTCAGSAGA TTGGCCATTCT CTTAGTCTTT GGCCATTCTA GTGATCCAAA GGCCGTAAATG 720
 TTCCCACTCT ACAAATATGT TGACATCAAC ACATTTGGCC TCTCTGCTGA TGACATACGT 780
 25 GGCACTCAGT CCGTGTATGG AGACCCAAA GAGAACCAC GCTTGCCAAA TCCTGACAAAT 840
 TCAGAACCAAG CTCTCTGTGA CCCCAATTG AGTTTGTATG CTGCTACTAC CGTGGGAAAT 900
 AAGATCTTTT TCTTCAAGAA CAGGTTCTTC TGGCTGAAGG TTTCTGAGAG ACCAAAGACC 960
 AGTGTTAAT TAAATTCTTC CTTATGGCCA ACCTTGCCAT CTGCAATTGA AGCTGCTTAT 1020
 GAAATTGAAG CCAGAAATCA AGTTTTTCTT TTTAAGATG ACAAATACTG GTTAATTAGC 1080
 30 AATTTAAGAC CAGAGCCAAA TTATCCCAAG AGCATAACAT CTTTGTGTTT TCCTAACTTT 1140
 GTGAAAAAAA TTGATGCAGC TGTTTTAAAC CCACGTTTT ATAGGACCTA CTTCTTTGTA 1200
 GATAACCACT ATTTGAGGAT TGATGAAAG AGACAGATGA TGGACCTGGT TTATCCCAAA 1260
 CTGATTACCA AGAATCTCCA AGGAATCGG CCTAAAATTG ATGCAGTCTT CTACTCTAAA 1320
 AACCAATACT ACTATTCTCT CCAAGGATCT AACCAATTG AATATGACTT CCTACTCCAA 1380
 35 CGTATCACCA AACACTGAA AAGCAATAGC TGGTTTGGTT GTTGAAATG GTGTAATTAA 1440
 TGGTTTTGTG TAGTTCACTT CAGCTTAATA AGTATTTATT GCATATTGCT TATGCTCTCA 1500
 GTGTACCACT ACTTAGATAT ATGTATCATA AAAATAAAAT CTGTAAACCA TAGGTAATGA 1560
 TTATATAAAA TACATAATAT TTTTCAATTT TGAAACTCT AATTGTCCAT TCTTGTCTGA 1620
 CTCTACTATT AAGTTTGAAA ATAGTTACCT TCAAGCAAG ATAATTCTAT TTGAAGCATG 1680
 40 CTCTGTAAGT TGCCTCTCAA CATCCTTGA CTGAGAAATT ATACTTACTT CTGGCATAAC 1740
 TAAATTAAG TATATATATT TTGGCTCAA TAAATTTG

Seq ID NO: 28 Protein sequence:
 Protein Accession #: Eos sequence

45 1 11 21 31 41 51
 MKFLILLIQ ATASGALPLN SSTSLKNNV LFGERYLEKF YGLEINKLPV TKMKYSQNLN 60
 KEKIQEMQHF LGLKVTGQLD TSTLEMMHAF RCGVPDVHVF REMPGGPVWR KHYITYIRIN 120
 50 YTPDMNREIV DYAIRKAFQV WSNVTPKFS KINTGMADIL VVFARGAHGD PHAFDGGKGI 180
 LAHAFPGSGG IGGDAHFEDF EFWTHSGGT NLPLTAVHEI GHSGLGLHSS DPKAVMFPTY 240
 KYVDINTFRL SADDIRIQS LYGDPKENQR LPNPDNSEPA LCDPNLSFDA VTTVGNKIFF 300
 FKDRFPWLKV SERPKTSVNL ISSLWPTLPS GIEAAEIEA RNQVFLFKDD KYWLISNLRP 360
 EPNYPSIHS FGPFNVFKKI DAAVFNPRFY RTYFPVDNQY WRYDERRQMM DPGYFKILITK 420
 55 NFQGIQPKID AVFYSKNKY YFFQGSNQFE YDFLLQRIK TLKSNWFGC

Seq ID NO: 29 DNA sequence
 Nucleic Acid Accession #: NM_006115.1
 Coding sequence: 236..1765

60 1 11 21 31 41 51
 GCTTCAGGGT ACAGCTCCCC CGCAGCCAGA AGCCGGGCTC GCAGCCCTC AGCACCCTC 60
 CGGGACACCC CACCCGCTTC CCAGGCGTGA CCTGTCAACA GCAACTTCGC GGTGTGGTGA 120
 65 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180
 GAGACCTAGA AATCCAAGCG TTGGAGGTCC TGAGGCCAGC CTAAGTCGCT TCAAAATGGA 240
 ACGAAGCGGT TTGTGGGGTT CCATTCAAGC CCGATACATC AGCATGAGTG TGTGACAAG 300
 CCGACGGAGA CTTGTGGAGC TGGCAGGGCA GAGCTGTCTG AAGGATGAGG CCCTGGCCAT 360
 70 TGCCGCGCTG GAGTGTCTGC CCAGGGAGCT CTTCCCGCCA CTCTTCATGG CAGCCTTTGA 420
 CGGGAGACAC AGCCAGACCC TGAAGGCAAT GGTGCAGGCC TGGCCCTTCA CCTGCCTCCC 480
 TCTGGGAGTG CTGATGAAGG GACAACATCT TCACCTGGAG ACCTTCAAAG CTGTGCTTGA 540
 TGGACTTGAT GTGCTCCTTG CCCAGGAGGT TCGCCCCAGG AGGTGGAAAC TTCAAGTGCT 600
 GGATTTACGG AAGAACTCTC ATCAGGACTT CTGGACTGTA TGGTCTGGAA ACAGGGCCAG 660
 75 TCTGTACTCA TTTCCAGAGC CAGAAGCAGC TCAGCCCATG ACAAAGAAGC GAAAAGTAGA 720
 TGGTTTGAGC ACAGAGGCAG AGCAGCCCTT CATTCCAGTA GAGGTGCTCG TAGACCTGTT 780
 CCTCAAGSAA GGTGCTCTGT ATGAATTGTT CTCTACCTC ATTGAGAAAG TGAAGCGAAA 840
 GAAAAATGTA CTAGCCCTGT GCTGTAAGAA GCTGAAGATT TTTGCAATGC CCATGCAGGA 900
 TATCAAGATG ATCCTGAAAA TGGTGCAGCT GGACTCTATT GAAGATTGGG AAGTGACTTG 960
 80 TACCTGGAAG CTACCCACTT TGGGAAAATT TTCTCCTTAC CTGGGCCAGA TGATTAACT 1020
 GCGTAGACTC CTCCTCTCCC ACATCCATGC ATCTTCTCTAC ATTTCCCGCG AGAAGGAAGA 1080
 GCAGTATATC GCCCAGTTCA CCTCTCAGTT CCTCAGTCTG CAGTGCCCTG AGGCTCTCTA 1140
 TGTGACTCTT TTTATTTTCC TTAGAGGCCG CCTGGATCAG TTGCTCAGGC ACGTGATGAA 1200
 CCCCTTGGAA ACCCTCTCAA TAACTAACTG CCGGCTTTG GAAGGGGATG TGATGCATCT 1260
 85 GTCCAGAGT CCAGCGTCA GTCAAGCTAAG TGTCTCAGT CTAAGTGGGG TCATGCTGAC 1320
 CGATGTAAGT CCGAGGCCCT TCCAAGCTCT GCTGGAGAGA GCCTCTGCCA CCCTCCAGGA 1380
 CCTGGTCTTT GATGAGTGTG GGATCACGGA TGATCAGCTC CTGCCCCCTC TGCCTTCCCT 1440
 GAGCCACTGC TCCAGCTTAA CAACCTTAAG CTTCTACGGG AATTCATCTT CCATATCTGC 1500

5

10

15

20

25

30

35

40

45

50

55

CTTGACAGT CTCCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCACG TGCTGTATCC 1560
 TGTCCCCCTG GAGAGTTATG AGGACATCCA TGGTACCCCTC CACCTGGAGA GGCTTGCCCTA 1620
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCACGCA TGGTCTGGCT 1680
 TAGTGCCAAAC CCTGTGCTCT ACTGTGGGGA CAGAACCTTC TATGACCCGG AGCCCATCCT 1740
 GTGCCCCCTG TTCAATGCCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTGGAAGCAA CAAGCAGGCC ACAGTTTCAG 1860
 ACAAAATGTT AGTGTGAGTG AGGAAAACAT GTTCAGTGAG GAAAAACAT TCAGACAAAT 1920
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTGACTTG AGGAGTTAAT 1980
 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAGAAAC 2100
 TGTGAAAAAT AAAGAGAAGC AATGTGAAGC AAAAAAATA AAAAAAATA

Seq ID NO: 30 Protein sequence:
 Protein Accession #: NP_006106.1

1 11 21 31 41 51
 | | | | |
 GCTTCAGGGT ACAGCTCCCC CGCAGCCAGA AGCCGGGCGCT GCAGCGCCCTC AGCACCGCTC 60
 CGGGACACCC CACCCGCTTC CCAGGCGTGA CCTGTCAACA GCAACTTCGC GGTGTGGTGA 120
 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180
 GAGACCTAGA AATCCAAGCG TTGGAGGTCC TGAGGCCAGC CTAAGTCGCT TCAAAATGGA 240
 ACGAAGGGCT TTGTGGGGTT CCATTCAAGC CGGATACATC AGCATGAGTG TGTGGACAAG 300
 CCCACGGAGA CTGTGGGAGC TGGCAGGGCA GAGCCTGCTG AAGGATGAGG CCCTGGCCAT 360
 TGCCCGCCCTG GAGTTGCTGC CCAGGGAGCT CTTCGCGCCA CTCTTCATGG CAGCCTTTGA 420
 CGGGAGACAG AGCCAGACCC TGAAGGCAAT GGTGCAGGCC TGGCCCTTCA CCTGCGCTCC 480
 TCTGGGAGTG CTGATGAAGG GACAACATCT TCACCTGGAG ACCTTCAAAG CTGTGCTTGA 540
 TGGACTTGAT GTGCTCCTTG CCCAGGAGGT TCGCCCCAGG AGGTGGAAAC TTCAAGTGCT 600
 GGATTTACGG AAGAAGCTCTC ATCAGGACTT CTGAGCTGTA TGGTCTGGAA ACAGGGCCAG 660
 TCTGTACTCA TTTCAGAGC CAGAAGCAGC TCAGCCCATG ACAAGAAGC GAAAAAGTAG 720
 TGGTTTGAGC ACAGAGGCGAG AGCAGCCCTT CATTCAGTA GAGGTGCTCG TAGACCTGTT 780
 CCTCAAGGAA GGTGCTCTGT ATGAATTTGT CTCTACCTC ATTGAGAAAG TGAAGCGAAA 840
 GAAAAATGTA CTACGCTGT GCTGTAAGAA GCTGAAGATT TTTGCAATGC CCATGCAGGA 900
 TATCAAGATG ATCTCGAAAA TGGTGACAGT GGAATCTATT GAAGATTGG AAGTGACTTG 960
 TACCTGGAAG CTACCCACCT TGGCGAAATT TTCTCCTTAC CTGGGCCAGA TGATTAATCT 1020
 GGTAGATGCT CTCTCTCCC ACATCCATGC ATCTTCTTAC ATTTCCCCGG AGAAGGAAGA 1080
 CAGATATATC GCGCAGTTGC CCTCTCAGT CCTCAGTCTG CAGTGCCTGC AGGCTCTCTA 1140
 TGTGACTTCT TTATTTTTC TTAGAGGCGC CCGGATCAG TTGCTCAGG ACGTGATGAA 1200
 CCCCTTGGAA ACCCTCTCAA TAACTAACTG CCGGCTTTCC GAAGGGGATG TGATGCATCT 1260
 GTCCAGAGT CCCAGCGTCA GTCAAGTAAG TGTCTGAGT CTAAGTGGGG TCATGCTGAC 1320
 CGATGTAAGT CCGAGGCCCT TCCAAGCTCT GCTGGAGAGA GCCTCTGCCA CCTCCAGGA 1380
 CCTGTGCTTT GATGAGTGTG GGATCAGCGA TGATCAGCTC CTGCCCCCTC TGCCCTCCCT 1440
 GAGCCACTGC TCCAGCTTAA CAACCTTAAG CTTCTACGGG AATTCATCT CCATATCTGC 1500
 CTGACAGAGT CTCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCACG TGCTGTATCC 1560
 TGTCCCCCTG GAGAGTTATG AGGACATCCA TGGTACCCCT CACCTGGAGA GGCTTGCCCTA 1620
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCACGCA TGGTCTGGCT 1680
 TAGTGCCAAAC CCTGTGCTCT ACTGTGGGGA CAGAACCTTC TATGACCCGG AGCCCATCCT 1740
 GTGCCCCCTG TTCAATGCCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTGGAAGCAA CAAGCAGGCC ACAGTTTCAG 1860
 ACAAAATGTT AGTGTGAGTG AGGAAAACAT GTTCAGTGAG GAAAAACAT TCAGACAAAT 1920
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTGACTTG AGGAGTTAAT 1980
 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAGAAAC 2100
 TGTGAAAAAT AAAGAGAAGC AATGTGAAGC AAAAAAATA AAAAAAATA

Seq ID NO: 31 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 64-2754

60

65

70

75

80

85

1 11 21 31 41 51
 | | | | |
 GGCAGGTCTC GCTCTCGGCA CCTCCCGGCG GCGCGCGTTC TCTGGGCGCT GCGCGGCATC 60
 CCGATGGCGG CCGCTGGGCG CCGCGGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120
 CTGACCCCTG TGATCTTCAG TCGTGATGTT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180
 CCTTCTAAAC TAGAGGCGCA CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240
 TCTGCAGACC TCATCGGCTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGGTCACTG 300
 TACACAGCCA GGGCTGTGTC GCTGTCTGAT AAGAAAAGAT CATTACCAT ATGGCTTTCT 360
 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420
 TCGAAGACAA GACACACTAG AGAAACCTGT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480
 ATTCTTGTCT CTATGCAAGA GAATTCCTTG GCGCCCTTCC CAITGTTTCT TCAACAAGTT 540
 GAATCTGATG CAGCAGAGAA CTATCTGTC TTCTACTCAA TAAGTGGAGG TGGAGTTGAT 600
 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG 660
 CCTGTGGATG TGAGAAGATA TGATGTTTTT GATTTGATTG CTTATGCTGC AACTGCAGAT 720
 GGATATTCAG CAGATCTGCC CCTCCCACTA CCCATCAGGG TAGAGGATGA AATGACAAC 780
 CACCCGTGTT TCACAGAAGC AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840
 ACTACAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGGCG 900
 CTGAATATCA GCATTTTGCA GCAGACACCA AGGTCACTTG GGTCTTTTTC TGTGCATCCC 960
 AGCAGAGCGG TAATCACACAC AGTCTCTCAT TATTTGGACA GAGAGGTTGT AGACAAGTAC 1020
 TCATTGATAA TGAAGTATCA AGACATGGAT GGCCAGTTTT TTGATTGAT AGGCACATCA 1080
 ACTTGATATA TAACGATAAC AGATTCAAAAT GATAATGCAC CCACCTTTCAG ACAAATGCT 1140
 TATGAAGCAT TTGTAGAGGA AATGCATTTC AATGTGGAAA TCTTACGAAT ACCTATAGAA 1200
 GATAAGGATT TAATTAAACAC TGCCAAATTG AGAGTCAATT TTACCAATTT AAAGGGAAAT 1260
 GAAAAATGAC ATTTCAAAAT CAGCAGAGAC AAAGAAACTA ATGAAGGTGT TCTTTCTGTT 1320
 GTAAAGCCAC TGAATTATGA AGAAAAAGGT CAAGTGAACC TGGAAATTTG AGTAAACAAT 1380
 GAAGCGCCAT TTCTAGAGCA TATTCACAGA GTGACAGCCT TGAACAGAGC CTTGGTTACA 1440
 GTTCATGTGA GGGATCTGGA TGAGGGGCGT GAATGCATCT CTGACGCCCA ATATGTGCGG 1500
 ATTAAGAAAT ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA 1560

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

AATAGAAATG GCAATGGTTT AAGGTACAAA AAATTGCAATG ATCCTAAAGG TTGGATCACC 1620
 ATTGATGAAA TTTCAGGGTC AATCATAACT TCCAAAATCC TGGATAGGGA GGTGAAACT 1680
 CCCAAAAATG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAGATGA TAGATCATGT 1740
 ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAAATGATA ATCCACCAGA AATACTTCAA 1800
 GAATATGTAG TCATTGCGAA ACCAAAAATG GGGTATACCG ACATTTTAGC TGTGTATCCT 1860
 GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC 1920
 AGTAGACTGT GGAGCCTCAC CAAAGTTAAT GATACAGCTG CCGGTCTTTC ATATCAGAAA 1980
 AATGCTGGAT TTCAAGAATA TACCATTCTC ATTACTGTAA AAGACAGGGC CGGCCAAGCT 2040
 GCAACAAAAT TATTGAGAGT TAATCTGTGT GAATGTACTC ATCCAACTCA GTGTGCTGG 2100
 ACTTCAAGGA GTACAGGAGT AATACTTGGG AAATGGGCAA TCCTTGCAAT ATTACTGGGT 2160
 ATAGCACTGC TCTTTCTGT ATTGCTAAT TTAGTATGTG GAGTTTTTGG TGCAACTAAA 2220
 GGGAAACGTT TTCTGAAGA TTTAGCACAG CAAAACCTAA TTATATCAAA CACAGAAGCA 2280
 CTTGGAGACG ATTAGGTGTG CTCTGCCAAT GGATTTATGA CCCAACTAC CAACAACCT 2340
 AGCCAAGGTT TTGTGGTAC TATGGGATCA GGAATGAAAA ATGGAGGGCA GGAAACCAAT 2400
 GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCCTGCC GGGGGGCTGG GCATCATCAT 2460
 ACCCTGGAGT CCTGAGGGGG AGGACACACG GAGGTGGACA ACTGCAGATA CACTTACTCG 2520
 GAGTGGCACA GTTTTACTCA ACCCGTCTC GGTGAAAAAT TGATCGATG TAATCAGAAT 2580
 GAAGACCCGA TGCCATCCCA AGATTATGTC CTCACITATA ACTATGAGGG AAGAGGATCT 2640
 CCAGCTGGTT CTGTGGGCTG CTGCAGTGA AAGCAGGAAG AAGATGGCT TGACTTTTTA 2700
 AATAATTGGG AACCCAAAT TATTACATTA GCAGAAAGCAT GCACAAAGAG ATAATGTCAC 2760
 AGTGCTACAA TTAGGTCTTT GTACAGACAT CTGGAGGTTT CCAAAAATAA TATTGTAAAG 2820
 TTCAATTTCA ACATGTAGT ATATGATGAT TTTTTTCTCA ATTTTGAATT ATGCTACTCA 2880
 CCAATTTATA TTTTAAAGC CAGTTGTGTC TTATCTTTTC CAAAAAGTGA AAAATGTTAA 2940
 AACAGACAAC TGGTAAATCT CAAACTCCAG CACTGGAATT AAGGTCTCTA AAGCATCTGC 3000
 TCTTTTTTTT TTTTACGGAT ATTTTAGTAA TAAATATGCT GGATAAATAT TAGTCCAACA 3060
 ATAGCTAAGT TAGCTTAATA TCACATTATT ATGTATTTCAC TTTAAGTGAT AGTTTAAAAA 3120
 ATAAACAAGA AATATTGAGT ATCACTATGT GAAGAAAGTT TTGGAAAAGA AACAATGAAG 3180
 ACTGAATTAA ATTAATAATG TTGCAGCTCA TAAAGAATTG GGACTCAACC CTACTGCAC 3240
 ACCAAATTC TTTGACTTTG GAGGCAAAAT GTGTTGAAGT GCCCTATGAA GTAGCAATTT 3300
 TCTATAGGAA TATAGTTGGA AATAAATGTG TGTGTGTATA TTATTATTA TCAATGCAAT 3360
 ATTTAAATAT AATAGAGAAC AAGAGGAAAA ATGGTAAAAA CTTGAAATGA GGCTGGGGTA 3420
 TAGTTTGTCC TACAATAGAA AAAAGAGAGA GCTTCTAGG CTTGGGCTCT TAAATGCTGC 3480
 ATTATACTG AGTCTATGAG GAAATAGTTC CTGTCCAATT TGTGTAATT GTTTAAAAAT 3540
 GTAATAAAT TAAACTTTTC TGGTTTCTGT GGGAAAGGAA TAGGGAATCC AATGGAACAG 3600
 TAGCTTTGCT TGTGAGTCTG TTTCAAGATT TCTGCATCCA CAAGTTAGTA GCAAACTGGG 3660
 GAATACTGCG TGCACTGGG GTTCCCTGCT TTTTGTAGC AAGGGTCCAG AGATGAGGTTG 3720
 TTTTTTCAAG GGAGCTAATA ACAAACAAT TTTAAACCT ACCTTACTG AAGTTAAATC 3780
 CTCTATTGCT GTTCTATTTC TCTCTTATAG TGACCAACAT CTTTTTAAT TAGATCCAAA 3840
 TAACCATGTC CTCTAGAGT TTAGAGGCTA GAGGAGGCTG AGGGGAGGAT CTTACTGAAA 3900
 GCACCTCGGG GAGATTGATT GTCTTAAAC CTAAGCCCA CAAACTTGAC ACCTGATCAG 3960
 GTCTGGGAGC TACAAAATTT CATTTTCTC CTCACCTGCC TTCTCTGAG TGGCATTTGC 4020
 CTGAATCAAG GAAAGCCAGG CTTGTGGGC CCCCCTCTT GGGCTTTCTG CTAAGCAAC 4080
 ACCTCCAGCA GAGATTCCCT TAASTGACTC CAGGTTTTCC ACCATCCTTC AGCGTGAAT 4140
 AATTTTTAAT CAGTTTGCTT TCTCCAGAGA AATTTTAAAA TAATAGAAGA AATAGAAAT 4200
 TTGAATGTAT AAAAGAAAAA GATCAAGTTG TCATTTTAGA ACAGAGGGAA CTTTGGGAGA 4260
 AAGCAGCCCA AGTAGGTTAT TTGTACAGTC AGAGGGCAAC AGGAAGATGC AGGCCTTCAA 4320
 GGGCAAGAG AGGCCACAGG GAATATGGGT GGGAGTAAAA GCAACATCGT CTGCTTATA 4380
 CTTTTTCTA GGCCTGGCAC TGCCTTTTCC TTTCTCAGGC CAATGGCAAC TGCCATTGGA 4440
 GTCCGGTGAG GGATCAGCCA ACCTCTTCTC TATGGCTCAC CTTATTGGA GTGAGAAATC 4500
 AAGGAGACAG AGCTGACTGC ATGATGAGTC TGAAGGCATT TGCAGGATGA GCCTGAAGT 4560
 GTTGTGAGA ACAAACAAGG CATTCTAGG AATTGTTGTA TTCTTCTGC AGCCCTCCTT 4620
 CTGGGCATA AGAAGGTCTA TGAATTAAAT GCTATCTAA AATCTGATT TATTCCTACA 4680
 TTTTCTGTT TCTAATTGGA CCCTAAAAAT TATGTGTTTT AGACTTAGAC TTTTATTGTC 4740
 CCCCCCCCC TTTTTTTTT AGACGGAGTC TCGCTCTGAC GCACAGGCTG GAGTGCAGTG 4800
 GCTCGGATCT CTGCTACTG AAGCTCCGC CTCCCGGTT CATGCCATT TCCTGCCTCA 4860
 GCTCCTGAG TAGCTGGGAC TACAGGCGCC CACCACCAG CCGGGCTAAT TTTTGTATT 4920
 TTTAATAGAG ACGSGGTTT ACTGTGTTAG CCAGGATGGT CTGATCTCC TGACCTCGTG 4980
 ATCCGCTGCT CTGGGCTCC CAAAGTGCTG GGATTACAGG CATGACCCAC CGCTCCCGGC 5040
 CTGTGTTTTG GTTTAAAGTC GTCTCTTTT AATGTAATCA TTTTGAACAT GTGTGAAAGT 5100
 TGATCATACG AATTGGATCA ACTTGAAAT ACTCAACCAA AAGACAGTCG AGAAGCCAGG 5160
 GGGAGAAAGA ACTCAGGCA CAAATATTG GTCTGAGAAAT GGAATTCTCT TGAAGCCTAG 5220
 TTGCTGAAAT TTCTGCTGT AACCAGAAGC CAGTTTTATC TAAAGGCTAC TGAACACACC 5280
 ACTGTGTTTT GCTCACTCCG TCACTCACC ATCAAAACCT GCTACCTCC CAAGACTTTA 5340
 CTAGTGCCGA TAACTTTCT CAAAGAGCAA CCAGTATCAC TTCCCTGTTT ATAAAACCTC 5400
 TAACCATCTC TTGTCTCTT GAACATGCTG AAAACCACT GGTCTGCATG TATGCCCGAA 5460
 TTTGTAATTC TTTTCTCTA AATGAAAT TAAATTTAGG GATTCAATTC TATATTTTCA 5520
 CATATGTAGT ATTATTATTT CCTTATATG GTAAGGTGAA ATTTATGGTA TTTGAGTGTG 5580
 CAAGAAATA TATTTTTAAA GCTTTTCAAT TTCCCCAGT GAATGATTTA GAATTTTTTA 5640
 TGTAAATATA CAGAATGTT TTTCTTACT TTATAAGGAA GCAGCTGTCT AAAATGCAGT 5700
 GGGGTTTGT TTGCAATGTT TTAACAGAG TTTTAGTATT GCTATTAAAA GAAGTTACTT 5760
 TGCTTTTAAA GAACTTGGC TGCTTAAAT AAGCAAAAT TGGATGCATA AAGTAATATT 5820
 TACAGATGTG GGGAGATGTA AAAAAACAAT ATTAACCTGG TTTCTGTTT TTGCTGTATT 5880
 TAGAGATTAA ATAATTCTAA GATGATCACT TTGCAAAAT ATGCTTATGG CTGGCATGGA 5940
 AATAGAAATA CTCATTTATG TCTTTGTTG ATTAATGGGG AATATTTTGG ACAATGTTT 6000
 ATTATCAAT TGTCGACATC ATTAATATAT ATTGTAATGT TGGGAAGAGA TCACTATTTT 6060
 GAAGCACAGC TTACAGATG AGTATCTATG ATACATATGT ATAATAAAT TTGATCGGGT 6120
 ATTAAGAGTA TTGAAGGTG GTTATAATG CAGAGTATTC CATGAATAGT ACATGACAC 6180
 AGGGGTTTTA CTTTGAAGC CAGTGTAGTC AAGGGAAGC ATGAGTTAAA AAGAAAGCA 6240
 GGCAATATTG CAGTCTGAT TCTGCCACT ACAGGATAGA TAATGCTGTA ACTTTAATGA 6300
 CAGATGATC CAACATGAA GGTGCTCTGT GCTTCAAGT GAATCTTTT CCATGACAGG 6360
 AGTGTGCTCC CTAACAAGC TTAAGACTGA TCATTTCAAA AATCTATTAG CTATATCAA 6420
 AGCCTTACAT TTTAATATAG GTTGAACCAA AATTTCAAT CCAGTAACCT CTATTGTAAC 6480
 CATTATTTTT GTGATTTCTG TCAAGAAATG TCATTGGATT TTTGTTGTA ATAGTAAAT 6540
 ACCGATATCA TTTCAGCTG CCTTCAGTAT TGATTTGGT GAATTTGGG TCATAATGTT 6600
 TGAGAGCAAT GGACATAGA GCCAGAATGC TTGGATATGA ATCCTGATC TGTCACTTAC 6660
 TTCTGTGTGA CTTTGAAGG GCTACTTAT TCTCTCTTA GCTTCTCAT TAAATCAAT 6720
 GAACATATGC AGCCTCATGG GGTGTTGAA TGATTAATAT AGTTAATATA CCTAAGATAC 6780

ATAGAACT GCCTGCACAT AGTAAAGAA TTATAAGTGT GAGGTAGTGT GTAAATTTAT 6840
 GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTT AGGGAAATAA AGTTTGTGCA 6900
 TATATATAAT CCGGAAACAT G

5 Seq ID NO: 32 Protein sequence:
 Protein Accession #: NP_001932.1

1	11	21	31	41	51	
MAAAGPRRSV	RGAVCLHLLL	TLVIFSRDGE	ACKKVILNVP	SKLEADKIIG	RVNLEECFRS	60
ADLIRSSDPD	FRVLNDGSVY	TARAVALSDK	KRSFTIWLSD	KRKQTQKEVT	VLLHQKQKVS	120
KTRHRETVL	RRAKRRWAPI	PCSMQENSIG	PFPLFLQQVE	SDAAQNYTVF	YSISGRGVDK	180
EPLNLFYIER	DTGNLFCRTP	VDREYDVDFD	LIAYASTADG	YSADLPLPLP	IRVEDENDNH	240
FVFTAIYNF	EVLESSRPGT	TVGVVCATDR	DEPDTMHTRL	KYSILQQTTPR	SPGLFSVHPS	300
TGVITTVSHY	LDREVVDKYS	LIMKVQMDMG	QFFGLIGTST	CIITVTDSDN	NAPTFRQNAV	360
EAFVEENAFN	VEILRIPIED	KDLINTANWR	VNFTILKQNE	NGHFKISTDK	ETNEGVLVSV	420
KPLNVEENRQ	NVLEIGVNNE	APFARDIPRV	TALNRALVTV	HVRDLDEGPE	CTPAAQYVRI	480
KENLAVGSKI	NGYKAYDPEN	RNGNGLRYKK	LHDPKGWITI	DEISGSIITS	KILDRVETP	540
KNELYNTVL	AIDKDRSCT	GTLAVNIEDV	NDNPPEILQE	VYVICKPKMG	YTDILAVDPD	600
EPVHGAPPYP	SLPNTSPZIS	RLWSLTKVND	TAARLSYQKN	AGFQEYTIPI	TVKDRAGQAA	660
TKLLRVNLCE	CTHPTQCRAT	SRSTGVILGK	WAILAILLGI	ALLPVSILLTL	VCGVFGATKG	720
KRFPEDLAQ	NLIISNTEAP	GDDRVCSSANG	FMTQTNNSS	QGPGTIMGSG	MKGNGQETIE	780
MMKGGNQITL	SCRAGGHHHT	LDSCRGGHTE	VDNCRYTYSR	WHSFTQPRLG	EKLHRCNQNE	840
DRMPSQDYL	TYNYBGRGSP	AGSVGCCSEK	QEEDGLDFLN	NLEPKFITLA	EACTKR	

Seq ID NO: 33 DNA sequence
 Nucleic Acid Accession #: Bos sequence
 Coding sequence: 64-2583

1	11	21	31	41	51	
GGCAGGTCTC	GCTCTGGCCA	CCCTCCCGGC	GCCCGCGTTC	TCCTGGCCCT	GCCCGGCATC	60
CCGATGGCCG	CCGCTGGGCC	CCGGCGCTCC	GTGCGCGGAG	CCGCTGCTCT	GCATCTGCTG	120
CTGACCTCTG	TGATCTTCAG	TGCTGATGGT	GAAGCCTGCA	AAAAGGTGAT	ACTTAATGTA	180
CTTCTTAAC	TAGAGGCAGA	CAAAATAATT	GGCAGAGTTA	ATTTGGAAGA	GTGCTTCAGG	240
TCGACAGACC	TCATCGGCTG	AAGTGATCCT	GATTTCAGAG	TTCTAAATGA	TGGGTCAGTG	300
TACACAGCCA	GGGCTGTGTC	GCTGTCTGAT	AAGAAAAGAT	CATTACCAT	ATGGCTTTCT	360
GACAAAAGGA	AACAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAGAAGGTA	420
TGGAAGACAA	GACACACTAG	AGAAACTGTT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCACCT	480
ATTCTTGTCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTCC	CATTGTTTCT	TCAACAAGTT	540
GAATCTGAT	CAGCAGACGA	CTATCTGTCT	TTCTACTCAA	TAACTGGAGC	TGGAGTTGAT	600
AAAGAACCTT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAAATCTATT	TTGCACTCGG	660
CCTGTGGATC	GTGAAGAAAT	TGATGTTTTT	GATTGTGATG	CTTATGCGTC	AACATGCAGAT	720
GGATATTCTG	CAGATCTGCC	CCTCCCACTA	CCCATCAGGG	TAGAGGATGA	AAATGACAAC	780
CACCTGTGTT	TCACAGAAGC	AAITTTATAAT	TTTGAAGTTT	TGGAAGTAG	TAGACCTGGT	840
ACTACAGTGG	GGGTGGTTTG	TGCCACAGAC	AGAGATGAAC	CGGACACAAT	GCATACCGCG	900
CTGAATACAC	GCATTTTGCA	GCAGACACCA	AGGTCACTGT	GGCTCTTTTC	TGTGCATCCC	960
AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
TCATTGATGA	TGAAGGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTGAT	AGGCACATCA	1080
ACTTGTATCA	TACAGTAAC	AGATTCAAAT	GATAATGCAC	CCACTTTTCA	ACAAATGCT	1140
TATGAAGCAT	TTGTAGAGGA	AAATGCATT	AATGTGGAAA	TCTTACGAAT	ACCTATAGAA	1200
GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTTT	AAAGGGAAAT	1260
GAAATGAGAC	ATTTCAAAAT	CAGCAGACAC	AAAGAAACTA	ATGAAGGTGT	TCTTTCTGTT	1320
GTAAAGCCAC	TTTCAGGGTG	AGAAAACCGT	CAAGTGAACC	TGGAATTTGG	AGTAAACAAT	1380
GAAGCGCCAT	TTGCTAGAGA	TATTCOCAGA	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
GTTCATGTGA	GGGATCTGGA	TGAGGGGCGT	GAATGCACCT	CTGCAGCCCA	ATATGTGCGG	1500
ATTAAAGAAA	ACTTAGACAGT	GGGTCAAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCAAT	ATCTTAAAGG	TTGGATCACC	1620
ATTGTATGAA	TTTCAGGGTG	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GTTTGAACCT	1680
CCCAAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAATGATA	ATCCACCAGA	AATACTTCAA	1800
GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCACTTTGC	CCAATACTTC	TCCAGAAATC	1920
AGTAGACTGT	GGAGCCTCAC	CAAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
AATGCTGGAT	TTCAAGAATA	TACCATTCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
GCAACAAAAT	TATGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACCTA	GTGCTGTGCG	2100
ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTTGCAAT	ATTACTGGGT	2160
ATAGCACTGC	TCCTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTGTGG	TGCAACTAAA	2220
GGGAAACGTT	TTCTCGAAGA	TTTAGCACAG	CAAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
CCTGGAGAGC	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAACTTAC	CAACAACCTC	2340
AGCCAAAGTT	TTTGTGGTAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAACCAAT	2400
GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCG	GGGGGCGTGG	GCATCATCAT	2460
ACCTGGAGCT	CCTGCAGGCG	AGGACACAGC	GAGGTGGACA	ACTGCAGATA	CACCTACTCG	2520
GAGTGGCACA	GTTTTACTCA	ACCCGCTCTC	GGTGAAGAAT	CCATTAGAGG	ACACACTGGT	2580
TAAAAATTAA	ACATAAAAGA	AATGTCATCG	ATGTAATCAG	AATGAAGACC	GCATGCCATC	2640
CCAGATTTAT	GTCTCACTTT	ATAACTATGA	GGGAAGAGGA	TCTCCAGCTG	GTTCGTGTGG	2700
CTGCTCAGT	GAAAGCAGG	AAGAAGATGG	CCTTGACTTT	TTAAATAAAT	TGGAACCCAA	2760
ATTATTATCA	TTAGCAGAAG	CATGCACAAA	GAGATAATGT	CACAGTGCTA	CAATTAGGTC	2820
TTTGTGCAGC	ATTCTGGAGG	TTTCCAAAAA	TAATATTGTA	AAGTTCAATT	TCAACATGTA	2880
TGTATATGAT	GATTTTCTTC	TCAATTTTGA	ATTATGCTAC	TCACCAATTT	ATATTTTATA	2940
AGCCAGTTGT	TGCTTATCTT	TTCCAAAAAG	TGAAAAATGT	TAAACAGAC	AACTGGTAAA	3000
TCACAACTC	CAGCACTGGA	ATTAAAGTCT	CTAAAGCATC	TGCTCTTTTT	TTTTTTTACG	3060
GATATTTTAG	TAAATAAAT	GCTGGATAAA	TATTAGTCCA	ACAATAGCTA	AGTTATGCTA	3120
ATATCACCATT	ATTATGTATT	CACTTTAAAT	GATAGTTTAA	AAAAATAACA	AGAAATATTG	3180
AGTATCATA	TGTGAAGAAA	GTTTGGGAAA	AGAAACAATG	AAGACTGAAT	TAAATTAATA	3240
ATGTTGCAGC	TCATAAGAAA	TTGGGACTCA	CCCTTACTGC	ACTACCAAT	TCATTTGACT	3300

5
10
15
20
25
30
35
40
45
50
55
60
65

TTGAGGCAA AATGTGTTGA AGTGCCTAT GAAGTAGCAA TTTTCTATAG GAATATAGTT 3360
GGAAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTAAAA ATGAAATGAG 3420
AACAAAGAGG AAAATGGTAA AACTTGAAA TGAGGCTGGG GTATAGTTTG TCCTACAATA 3480
GAAAAAAGAG AGAGCTTCCT AGGCTGGGC TCTTAAATGC TGCATTATAA CTGAGTCTAT 3540
GAGGAAATAG TTCTGTGCTA ATTTGTGTAA TTTGTTTAAA ATTGTAATAA AATTAAACTT 3600
TTCTGGTTTC TGTGGGAAGG AAATAGGGAA TCCAATGGAA CAGTAGCTTT GCTTTGCAGT 3660
CTGTTTCAAG ATTGTCTGAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT 3720
GGGGTTCCTT GCTTTTGTGT AGCAAGGGTC CAGAGATGAG GTGTTTTTTT CGGGGAGCTA 3780
ATAACAAAAA CATTTTAAAA CTTACCTTTA CTGAAGTTAA ATCCTCTATT GCTGTTTCTA 3840
TTCTCTCTTA TAGTGACCAA CATCTTTTAA ATTTAGATCC AAATAACCAT GTCTCTCTAG 3900
AGTTTAGAGG CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAAGCACCCT GGGGAGATTG 3960
ATTGTCTAAG TTGTCTAAGC CCACAACCTT GACACCTGAT CAGGTCCTGG AGCTACAAAA 4020
TTTCATTTTT CTCTCAGTCT CCCTTCTTCT GAGTGGCATT GGCTGAAATC AAGGAAAGCC 4080
AGGCTTGTGT GGGCCCCCTT TTTGGGCTTT CTGCTAAAGC AACACCTCCA GCAGAGATTG 4140
CCTTAAGTGA CTCCAGGTTT TCCACCATCC TTCAGGCTGA ATTAATTTTT AATCAGTTTG 4200
CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTTGAATG TATAAAAGAA 4260
AAGATCAAG TTGTCTAAGT AGAACAGAGG GAACCTTGGG AGAAAGCAGC CCAAGTAGGT 4320
TATTTGTACA GTGAGAGGGC AACAGGAAGA TGCAGGCCTT CAAGGGCAAG GAGAGGCCAC 4380
AAGGAATATG GGTGGGAGTA AAGCAACAT CGTCTGCTTC ATACTTTTTT CTAGGCTTGG 4440
CACTGCCTTT TCCTTTCTCA GGCCAAATGGC AACTGCCATT TGAGTCCGGT GAGGGATCAG 4500
CCAACCTCTT CTCTATGGCT CACCTTATTT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC 4560
TGCAATAGA GTCTGAAGCT ATTTGCAGGA TGAGCCTGAA CTGGTTGTGC AGAACAAAA 4620
AGGCATTATC GGGATTTGTT GTATTCTTTC TGCAGCCCTC CTCTGGGACA CTAAGAGGT 4680
CTATGAATTA AATGCTATC TAAAATTCTG ATTTATTCCT ACATTTTCTG TTTTCTAATT 4740
TGACCCCTAA ATCTATGTGT TTTAGACTTA GACTTTTTAT TGCCCCCCCC CCCTTTTTTT 4800
TTGAGACGGA GTCTGCTCTT GACGCACAGG CTGGAGTGCA GTGGCTCCGA TCTCTGCTCA 4860
CTGAAGCTCA CGCCTCCCGG GTTCATGCCA TTCTCTGCTC TCAGCTCTCT GAGTAGCTGG 4920
GACTACAGGC GCCCACCACC AGCCCGGCTT AATTTTTTGT ATTTTAAATA GAGACGGGGT 4980
TTCACTGTGT TAGCCAGGAT GGTCTCGATC TCCTGACCTC GTGATCCGCC TGCTCCGGCC 5040
TCCCAAAGTG AGGCATGACC CACCGCTCCC GGCTTGTGTT TCCGTTTAAA 5100
GTGCTCTTCT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTTGATCAT ACGAATTTGA 5160
TCAATCTTGA AATCTCTAAC CAAAAGACAG TCGAGAAGCC AGGGGGAGAA AGAACTCAGG 5220
GCACAAAAAT TTGCTCTGAG AATGGAATTC TCTGTAAGCC TAGTTGCTGA AATTCTCTGC 5280
TGTAACCAGA AGCCAGTTTT ATCTAACGGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340
CCCACTCACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTCCGC ATAAACTTTC 5400
TCAAAGAGCA ACCAGTATCA CTCCCTGTGT TATAAAACCT CTAACCATCT CTTTGTCTCT 5460
TGAACTAGCT GAAACCCACC TGGTCTGAT GTATGCCCGA ATTTGTAATT CTTTCTCTC 5520
AAATGAAAAA TTAATTTTAG GGATTCATTT CTATATTTTC ACATATGTAG TATTATTATT 5580
TCCTTATATG TGTAAAGTGA AATTTATGTT ATTTGAGTGT GCAAGAAAAA ATATTTTAAA 5640
AGCTTTCTAT TTTTCCCCAG TGAATGATT AGAATTTTTT ATGTAAATAT ACAGAATGTT 5700
TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAATGTCAG TGGGGTTTGT TTTGCAATGT 5760
TTTAAACAGA GTTTTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAACTTGG 5820
CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAAGTAATAT TTACAGATGT GGGGAGATGT 5880
AATAAAACAA TATTAACCTG GCTGCTTAAA ATAAGCAAAA ATTGGATGCA TAAAGTAATA 5940
TTTACAGATG TGGGGAGATG TAATAAAACA ATATTAACTT GGTTCCTTGT TTTTGTGTA 6000
TTTAGAGATT AAATAATTCT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCTGGCATG 6060
GAAATAGAAA TACTCAATTA TGTCTTTGTT GTATTAAATG GGAATATTTT GGACAATGTT 6120
TCATTATCAA ATTTGTCGAC TCATTAAAT ATATTGTAAT GTTGGGAAGA GATCACTATT 6180
TTGAAGCACA GCTTTACAGA TGAGTATCTA TGATACATAT GTATAATAAA TTTTGATCGG 6240
GTATTAAAAA TTTTGAAGAG TGGTTATAAT TGCAGAGTAT TCCATGAATA GTACACTGAC 6300
ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAA ACATGAGTTA AAAAGAAAAA 6360
CAGGCAATAT TGCAGTCTTG ATCTGCCAC TTACAGGATA GATAATGCCCT GAACCTTAAT 6420
GACAAGATGA TCCAACCAT AAGGTGCTCT GTGCTTACA GTGAATCTTT TCCCATGCA 6480
GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTC AATATCTATT AGCTATATCA 6540
AAAGCCTTAC ATTTTAAAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600
ACCATTATTT TTGTGTATGT CTTCAGAAAT GTTCATTGGA TTTTGTGTTG TAATAGTAAA 6660
ATACCGGATA CATTTCAAGT GTCTTCAGT ATTGATTGG TTGAATATTG GGTCAATAG 6720
GTTGAGAAGC ATGGACACTA GAGCCAGAAAT GCTTGGATAT GAATCTGGA TCTGTCACTT 6780
ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCT TAGCTTCTC ATTAATAATCA 6840
ATGAACAATG CCAGCTCAT GGGGTTGTTG AATGATTAAA TTAGTTAATA TACCTAAAGT 6900
ACATAGAACA CTGCTGCAC ATAGTAAAG AATTATAAGT GTGAGGTAGT TGGTAAAT 6960
ATGTAGTTGG ATATACCTACC GAACAATATC TAATCTCTTT TTAGGGAAT AAAGTTTGTG 7020
CATATATATA ATCCGAAAC ATG

Seq ID NO: 34 Protein sequence:
Protein Accession #: NP_077741.1

70
75
80
85

1 11 21 31 41 51
MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60
ADLIRSSDPD FRVLNDGSVY TARAVALSDK KRSFTIWLSD KRKQTQKEVT VLLEHQKKVS 120
KTRHRTRETV RRAKRWAPI PCSMQENSLG FFPLFLQQVE SDAQNYTVF YSISGRGVDK 180
EPLNLFYIER DTGNLFCFTRP VDREYDVFD LIAYASTADG YSADLPLPLF IRVEDENDNH 240
PVFTEAIYNF EVLESSRPQT TVGVVCATDR DEPDIMHTRL KYSILQQTTPR SPGLFSVHPS 300
TGVITTVSHY LDREVVVKYS LIMKVQDMDG QFFPLIGTST CIITVDSND NAPTFRQNAV 360
EAFVEENAFN VEILRIPIED KDLINTANWR VNPTILKNGE NGHFKISTDK ETNEGVLSSV 420
KPLVYENRQV VNLIGVNNV AFFARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI 480
KENLAVGSKI NGKYAYDPEN RENGNGLYKK LHDPKGWITI DEISGSIITS KILDRREVETP 540
KNELYNITVL AIDKDDRSCT GTLAVNIEDV NDNPPPEILQE YVVICPKMG YTDILAVDPD 600
EPVHGAPPYF SLENTSPEIS RLWSLTKVND TAARLSYQKN AGFQETIPI TVKDRAGQAA 660
TKLLRVNLCE CTFTPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720
KRFPEDLAQQ NLIISNTEAP GDRVCSANG FMTQTNNSS QGFCGTMSSG MKNGGQETIE 780
MMKGNQZLE SCRGAGHHT LDSRCGGHTE VDNCRITYSE NBSFTQPRIG EESIRGHTG

Seq ID NO: 35 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 146-1273.

1 11 21 31 41 51
5 GGGAGTGGGC GTGGCGGTGC TGCCAGGTG AGCCACCGCT GCTTCTGCCC AGACACGGTC 60
GCTCCACAT CCAGGTCTTT GTGCTCCTCG CTGCTCTGTT CCTTTTCCAC GCATTTTCCA 120
GGATAACTGT GACTCCAGGC CCGCAATGGA TGCCCTGCAA CTAGCAAAAT CGGCTTTTGC 180
CGTTGATCTG TTCAAACAAC TATGTGAAAA GGAGCCACTG GGCATATGTC TCTTCTCTCC 240
AATCTGTCTC TCCACCTCTC TGTCATTTGC TCAAGTGGGT GCTAAAGGTG ACATGCAAAA 300
TGAAATTGGA CAGGTTCTTC ATTTGAAAA TGTCAAAGAT ATACCCCTTG GATTTCAAAC 360
10 AGTAACATCG GATGTAAACA AACTTAGTTC CTTTACTCA CTGAACTPA TCAAGCGGCT 420
CTACGTAGAC AAATCTCTGA ATCTTTCTAC AGAGTTCATC AGCTCTACGA AGAGACCCCTA 480
TGCAAGAGAA TTGAAACTG TTGACTTCAA AGATAAATG GAAGAAACGA AAGGTGAGAT 540
CAACAATCA ATTAAGGATC TCACAGATGG CCACTTTGAG AACATTTTAG CTGACACAG 600
TGTGAACGAC CAGACCAAAA TCCTTGTGGT TAATGCTGCC TACTTTGTTG GCAAGTGGAT 660
15 GAAGAAATTT CCTGAATCAG AAACAAAGA ATGTCTTTC AGACTCAACA AGACAGACAC 720
CAAAACAGTG CAGATGATGA ACATGGAGGC CACGTTCTGT ATGGGAAACA TTGACAGTAT 780
CAATTGTAG ATCATAGAGC TTCTTTTCA AAATAAGCAT CTCAGCATGT TCATCTTACT 840
ACCCAAGAT GTGGAGGATG AGTCCACAGG CTTGGAGAAG ATTGAAAAAC AACTCAACTC 900
AGAGTCACTG TCACAGTGA CTAATCCAG CACCATGGCC AATGCCAAG TCAAACTCTC 960
20 CATTCCAAAA TTTAAGGTGG AAAAGATGAT TGATCCCAAG GCTTGTCTGG AAAATCTAGG 1020
GCTGAAACAT ATCTTCAGTG AAGACACATC TGATTTCTCT GGAATGTCAG AGACCAAGGG 1080
AGTGGCCCTA TCAAAATGTTA TCCACAAAGT GTGCTTAGAA ATAACGTGAAG ATGGTGGGGA 1140
TTCCATAGAG GTGCCAGGAG CACGGATCCT GCAGCACAA GATGAATTGA ATGCTGACCA 1200
TCCCTTTATT TACATCATCA GGCACAACAA AACTCGAAAC ATCATTTTCT TTGGCAAAAT 1260
25 CTGTTCTCCT TAAGTGGCAT AGCCCATGTT AAGTCTCTCC TGACTTTTCT GTGGATGCCG 1320
ATTTCTGTAA ACTCTGCATC CAGAGATTCA TTTTCTAGAT ACAATAAAT GCTAATGTTG 1380
CTGGATCAGG AAGCCGCGAG TACTTGTCT ATGTAGCCTT CACACAGATA GACCTTTTTT 1440
TTTTTCCAA TCTATCTTTT GTTTCTTTT TTTCCATAAG ACAATGACAT ACGCTTTTAA 1500
TGAAAAGGAA TCACGTTAGA GGA AAAATAT TTATTCATTA TTTGTCAAAT TGTCCGGGGT 1560
30 AGTTGGCAGA AATACAGTCT TCCACAAAGA AAATTCCTAT AAGGAAGATT TGAAGCTCT 1620
TCTTCCAGC ACTATGCTTT CCTTCTTTG GATAGAGAAT GTTCCAGACA TTCTCGTCTC 1680
CCTGAAAGAC TGAAGAAGT GTAGTGCATG GGACCCACGA AACTGCCCTG GCTCCAGTGA 1740
AACTGGGCA CATGCTCAGG CTACTATAGG TCCAGAAATC CTTATGTTAA GCCCTGGCAG 1800
GCAGGTGTTT ATTA AAAATTC TGAATTTTGG GGATTTTCAA AAGATAATAT TTTACATACA 1860
35 CTGTATGTTA TAGAATCTCA TGGATCAGAT CTGGGGCAGC AACCTATAAA TCAACACCTT 1920
AATATGCTGC AACAAAATG AGAATATTCA GACAAAATGG ATACATAAAG ACTAAGTAGC 1980
CCATAAGGGG TCAAAATTTG CTGCCAAATG CGTATGCCAC CAACCTACAA AAACACTTCG 2040
TTCCAGAGC TTTTCAGATT GTGGAATGTT GGATAAGGAA TTATAGACCT CTAGTAGCTG 2100
40 AAATGACAGA CCCCAAGAGG AAGTTCAGAT CTTAATATAA ATTCACCTTC ATTTTGTATA 2160
GCTGTCCCAT CTGCTCATGT GGTGGCACT AGACTGGTGG CAGGGGCTTC TAGCTGACTC 2220
GCACAGGGAT TCTCAATAA GCCGATATCA GAATTTGTGT TGAAGGAACT TGTCTCTTCA 2280
TCTAATATGA TAGCGGAAA AGGAGAGGAA ACTACTGCCT TTAGAAAATA TAAGTAAAGT 2340
GATTAAAGTG CTCACGTTAC CTTGACACAT AGTTTTTCAG TCTATGGGTT TAGTTACTTT 2400
45 AGATGGCAAG CATGTAACTT ATATTAAATAG TAATTTGTAA AGTTGGGTGG ATAAGCTATC 2460
CCTGTGGCCG GTTCACTGGAT TACTTCTCTA TAAAAAATAT ATATTACCA AAAAATTTTG 2520
TGACATTCCT TCTCCCATCT CTTCCTTGAC ATGCATTGTA AATAGGTCTT TCTTGTCTG 2580
AGATTCAATA TTGAATTTCT CCTATGCTAT TGACAATAAA ATATTATTGA ACTACC

Seq ID NO: 36 Protein sequence:
Protein Accession #: NP_002630.1

1 11 21 31 41 51
55 MDALQLANSA FAVDLFKQLC EKEPLGNVLF SPICLSTSLA LAQVGARGDT ANEIGQVLHF 60
ENVKDIIPFG QVTTSVNVKL SSFYSKLKIK RLYVDKSLNL STEFISSTKR PYAKELETVD 120
FKDKLEETKG QINNSIKDLT DGHFENILAD NSVNDQTKIL VVNAAYFVGK WMKFPESST 180
KECPFRLNKT DTKFVQMMNM EATPCMGNID SINCKIIELP PQNKLHSMFI LLPKDOVEDS 240
TGLEKIEKQL NSESLSQWYN PSTMANAKVK LSIPIKFKVEK MIDPKACLEN LGLKHIPSED 300
60 TSDFSMSSET KGVALSNIHV KVCLEITEDG GDSIEVPGAR ILQHKDELNA DHPPIYIIRH 360
NKTRNIIFPG KPCSP

Seq ID NO: 37 DNA sequence
Nucleic Acid Accession #: NM_0168583
Coding sequence: 72-842

1 11 21 31 41 51
70 GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60
TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCGAGA 120
CCATGGCCCA GTTTGAGGAG CTGCCCGTGC CCCTGGACCA GACCTGCCCC TTGAATGTGA 180
ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GCCCTCAGCA 240
ATGGCCTGCT GTCTGGGGGC CTGTTGGGCA TTCTGGAAAA CCTTCCGCTC CTGACATCC 300
75 TGAAGCCTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGGG AAAGTGACGT 360
CAGTGATTCC TGGCCTGAAC AACATCATTG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420
AACTTGGCCT TGTGACAGAG CCTGATGGCC ACCGTCTCTA TGTCAACATC CCTCTCGGCA 480
TAAAGCTCCA AGTGAATAGC CCCTGGTGG GTGCAAGTCT GTTGAGGCTG GCTGTGAAGC 540
TGGACATCAC TGGAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC 600
TTGGTGACTG CACCCATTCC CCTGGAAGCC TGCAAAATTC TCTGCTTGAT GGACTTGGCC 660
80 CCCTCCCATC TCAAGGTCTT CTGGACAGCC TCACAGGGAT CTTGAATAAA GTCCCTGCTG 720
AGTTGGTTCA GGGCAACGTG TGCCCTCTGG TCAATGAGGT TCTCAGAGGC TTGGACATCA 780
CCCTGGTGCA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTG ATCAAGGTCT 840
AAGCCTTCCA GGAAGGGGCT GGCCCTGTCT GAGCTGCTTC CCAGTGCTCA CAGATGGCTG 900
GCCCATGTGC TGGAGATGA CACAGTTGCC TTCTCTCCGA GGAACCTGCC CCCTCTCCTT 960
85 TCCACCAAGC GGTGTGTAAC ATCCATGTG CCTCACCTAA TAAATGGCT CTTCTTCTGC 1020
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 38 Protein sequence:
Protein Accession #: NP_057667

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      MFQTGGLIVF YGLLAQTMQA FGLPLVPLDQ TLPLNVNPAL PLSPTGLAGS LTNALSNGLL 60
      SGLLGLILEN LPLLDILKPG GGTSGGLLGG LLGKVTSPVP GLNNI IDIKV TDPQLLELGL 120
      VQSPDGHRLY VTIPLGILKI VNTPLVGASL LRLAVKLDIT AEILAVDRDQ ERIHLVLGDC 180
      THSPGSLQIS LLDGLGFLPI QGLDLSLTGI LNKVLPVLVQ GNVCPVLNVEV LRGLDITLVH 240
10     DIVNMLIHGL QFVIKV

```

Seq ID NO: 39 DNA sequence
Nucleic Acid Accession #: NM_004363.1
Coding sequence: 115-2223

```

15      1      11      21      31      41      51
      |      |      |      |      |      |
      CTCAGGCGAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 60
      TCCTGGAAGT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGAG 120
      TCTCCCTCGG CCCCTCCCA CAGATGGTGC ATCCCTCGGC AGAGGCTCCT GCTCAGAGCC 180
      TCACTTCTAA CCTTCTGGAA CCGGCCACCC ACTGCCAAGC TCACATATGA ATCCACGCCG 240
      TTCAATGTGG CAGAGGGGAA GGAGGTGCTT CTACTTGTC ACAAATCTGC CCAGCATCTT 300
      TTTGGCTACA GCTGGTACAA AGGTGAAAGA GTGGATGGCA ACCGTCAAAT TATAGGATAT 360
      GTAATAGGAA TCACAACAGC TACCCAGGCG CCGCATACA GTGGTGGAGA GATAATATAC 420
      CCCAATGATG CCTGCTGAT CCAGAACATC ATCCAGAATG ACACAGGATT CTACACCCCTA 480
      CACGTCAATA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCGG GGTATACCCG 540
      GAGCTGCCCA AGCCCTCCAT CTCCAGCAAC AACTCCAAC CCGTGGAGGA CAAGGATGCT 600
      GTGGCCTTCA CCTGTGAACC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAAT 660
      CAGAGCCTCC CGGTGAGTCC CAGGCTGCAG CTGTCCAATG GCAACAGGAC CCTCACTCTA 720
      TTCAATGTCA CAAGAAATGA CACAGCAAGC TACAAATGTG AAACCCAGAA CCCAGTGAGT 780
      GCGAGGCGCA GTGATTCAGT CATCTGAAT GTCCCTATG GCCCGGATGC CCCCAACATT 840
      TCCCTCTTAA ACACATCTTA CAGATCAGGG GAAATCTGA ACCTCTCCTG CCACGCAGCC 900
      TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTTCCAGCA ATCCACCCAA 960
      GAGCTCTTTA TCCCCCAACAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT 1020
      AACTCAGACA CTGGCCTCAA TAGGACCAAC GTCAACGACA TCACAGTCTA TGCAGAGCCA 1080
      CCAAAACCCCT TCATCACCA CACAACCTCC AACCCTGGG AGGATGAGGA TGCTGTAGCC 1140
      TTAACCTGTG AACTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC 1200
      CTCCCGGTCA GTCCAGGCT GCAGCTGTCC AATGACAACA GGACCTCAC TCTACTCAGT 1260
      GTCAACAAGG ATGATGTAGG ACCCTATGAG TGTGGAATCC AGAAGCAATT AAGTGTGAC 1320
      CACAGCGCAG CAGTATCTCT GAATGTCTCT TATGGCCCGA ACACCCACCA CATTTCCCCC 1380
      TCATACACCT ATTACCGTAT AGGGGTGAAC CTCAGCCTCT CTGCCATGC AGCCTCTAAC 1440
      CCACCTGCAC AGTATTCTTG GCTGATTGAT GGAACATCC AGCAACACAC ACAAGAGCTC 1500
      TTTATCTCCA ACATCACTGA GAAGAACAGC GGAATCTATA CCTGCCAGGC CAATAACTCA 1560
      GGCAGTGGCC ACAGCAGGAC TACAGTCAAG ACAATCACAG TCTCTGCGGA GCTGCCCAAG 1620
      CCTCCATCTC CCAGCAACAA CTCCAAACCC GTGGAGGACA AGGATGCTGT GGCCTTCACC 1680
      TGTGAACCTG AGGCTCAGAA CACAACCTAC CTGTGGTGGG TAAATGGTCA GAGCCTCCCA 1740
      GTCAGTCCCA GCGTGCAGCT GTCCAATGGC AACAGGACCC TCACTCTATT CAATGTCA 1800
      AGAAATGAGG CAAGAGCCTA TGTATGTGGA ATCCAGAACT CAGTGAGTGC AAACCCGAGT 1860
      GACCCAGTCA CCGTGGATGT CCTCTATGGG CCGGACACCC CCATCATTTT CCCCCAGAC 1920
      TGTCTTACCT TTTGGGAGGC GAACCTCAAC CTCTCCTGCC ACTGGGCTCT TAACCCATCC 1980
      CCGCAGTATT CTTGGCGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTATC 2040
      GCCAAATACA CGCCAAATAA TAACGGGACC TATGCTGTGT TTGTCTTAA CTTGGCTACT 2100
      GGCOCGCAAT ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TTCTCCTGGT 2160
      CTCTCAGCTG GGGCCATCAT CGGCATCATG ATTGAGTGTG TGGTGGGGT TGCTCTGATA 2220
      TAGCAGCCTT GGTGTAGTTT CTTCATTTCA GGAAGACTGA CAGTTGTTTT GCTTCTCTCT 2280
      TAAAGCATTT GCAACAGCTA CAGTCTAAAA TTGCTTCTTT ACCAAGGATA TTTACAGAAA 2340
      AGACTCTGAC CAGAGATCGA GACCATCCTA GCCAACATCG TGAAACCCCA TCTCTACTAA 2400
      AAATACAAAA ATGAGCTGGG CTTGGTGGCG CGCACCTGTA GTCCAGTTA CTCGGGAGGC 2460
      TGAGGAGGAG GAATGCTCTG AACCCTGGAG GTGGAGATTG CAGTGAGCCC AGATCGCACC 2520
      ACTGCATCTC AGTCTGGCAA CAGAGCAAGA CTCCATCTCA AAAAGAAAAG AAAAGAGAC 2580
      TCTGACCTGT ACTCTTGAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTTCCAA 2640
      AACTTTAATG AACTTAATGA CAGCTTCATG AAATGTGCCA CCAAGATCAA GCAGAGAAAA 2700
      TAATTAATTT CATGGGACTA AATGAACATA TGAGGATTGC TGATTTCTTA AATGTCTTGT 2760
      TTCCAGATT TCAGGAAACT TTTTCTCTTT TAAGCTATCC ACTCTTACAG CAATTTGATA 2820
      AATATACATT TTGTGAACAA AATTGAGAC ATTTACATTT TCTCCTATG TGGTCGCTCC 2880
      AGACTTGGGA AACTATTCAT GAATATTTAT ATTGTATGGT AATATAGTTA TTGCACAAGT 2940
      TCAATAAAAA TCTGCTCTTT GTATAACAGA AAAA

```

Seq ID NO: 40 Protein sequence:
Protein Accession #: NP_004354.1

```

75      1      11      21      31      41      51
      |      |      |      |      |      |
      MESPSAPPER WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEKKE VLLLVENLPQ 60
      HLPFGYSWKY ERVDGNRII GYVIGTQAT PGPAYSGREI IYPNASLLIQ NI IQNDTGFY 120
      TLHVIKSDLV NEBATGQFRV YPELPKPSIS SNNSKPFVEDK DAVAFCTEPE TQDATYLVWV 180
      NNQSLVSPR LQLSNGNRTL TLFNVTRNDT ASYKCTQNP VSARRSDSVI LNVLYGPDAP 240
      TISPLNTSYR SGENLNLSCH AASNPPAQYS WFNVTGPOQS TQELFIPNIT VNNSGSYTCQ 300
      AENSDTGLNR TTVTTITVVA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ NTVLWVWVN 360
      QSLFVSPRLQ LSNDRNTLTL LSVTRNDVGF YECGIONELS VDHSDPVILN VLYGDDFTI 420
      SPSYTYFRPG VNLSSLCHAA SNPPAQYSWL IDGNIQHTQ ELFISNITER NSGLYTCQAN 480
      NSASGHSRTT VKTITVSAEL FKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLWVWNGQS 540
      LPSVSPRLQLS NGNRLTLTFN VTRNDARAYV CGIQNSVSAN RSDPVTLDEL YGPDPTIISP 600
      PDSSYLSGAN LNLSCHSASN PBPQYSWRIN GIPQHTQVL FIAKITPNNN GTYACFVSNL 660
      ATGRNNSIVK SITVSASGTS PGLSAGATVG IMIGVLVGVA LI

```

Seq ID NO: 41 DNA sequence
Nucleic Acid Accession #: NM_006952.1
Coding sequence: 11-793

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
AATCCCGACA	ATGGCGAAAG	ACAACCTCAAC	TGTTGGTTC	TTCCAGGGCC	TGCTGATTTT	60
TGGAATGTG	ATTATTGGT	GTGGGGCAT	TGCCCTGACT	GCGGAGTGCA	TCTTCTTTGT	120
ATCTGACCAA	CACAGCCTCT	ACCCACTGCT	TGAAGCCACC	GACAACGATG	ACATCTATGG	180
GGCTGCCTGG	ATCGGCATAT	TTGTGGGCAT	CTGCCTCTTC	TGCTCTGCTG	TCTTAGGCAT	240
TGTAGGCATC	ATGAAGTCCA	GCAGGAAAT	TCTTCTGGCG	TATTTTCATT	TGATGTTTAT	300
AGTATATGCC	TTTGAAGTGG	CATCTTGTAT	CACAGCAGCA	ACACAACGAG	ACTTTTTCAC	360
ACCCAACCTC	TTCTGAAGC	AGATGCTAGA	GAGGTACCAA	AACAACAGCC	CTCCAACAA	420
TGATGACCAG	TGGAATAACA	ATGGAGTCAC	CAAAACCTGG	GACAGGCTCA	TGCTCCAGGA	480
CAATTGCTGT	GGCGTAAATG	GTCCATCAGA	CTGGCAAAAA	TACACATCTG	CCTTCCGGAC	540
TGAGAATAAT	GATGCTGACT	ATCCCTGGCC	TOGTCATGCG	TGTGTTATGA	ACAATCTTAA	600
AGAACCTCTC	AACTGGAGG	CTGTAAACT	AGGCGTGCCT	GGTTTTTATC	ACAATCAGGG	660
CTGCTATGAA	CTGATCTCTG	GTCCAAATGA	CCGACACGCC	TGGGGGGTTG	CCTGGTTTGG	720
ATTGGCATT	CTCTGCTGGA	CTTTTGGGT	TCTCTGGGT	ACCATGTTCT	ACTGGAGCAG	780
AATTGAATAT	TAAGAA					

Seq ID NO: 42 Protein sequence:
Protein Accession #: NP_008883.1

25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
MAKDNSTVRC	FOGLLIFGVN	IIGCCGIALT	AECIFVSDQ	HSLYPLLEAT	DNDDIYGAAW	60
IGIFVIGICLF	CLSVLGIVGI	MKSSRKILLA	YFILMPIVYA	FEVASCITAA	TQRDFPTPNL	120
FLKQMLERYQ	NNSPPNDDQ	WKNNGVTKTW	DRLMLQDNCC	GVNGPSDWQK	YTSAPRTENN	180
DADYWPWRQC	CVNNLKEPL	NLEACKLGP	GFYHNQGCYE	LISGPMNRHA	WGVANFGFAI	240
LCWTFWVLLG	TMFYWSRIEY					

Seq ID NO: 43 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 83-2605

40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
GCGGACAGCA	TCTGCGCGTA	TCCTGGAGCC	GGCCAGTTG	TGAAGTAGGA	GAGCTTTGGG	60
ACCTCTGTCC	CAAGCAAGAG	AGATGAATGG	AGAGTATAGA	GCGAGAGGAT	TTGGACGAGG	120
AAGATTTCAA	AGCTGGAAAA	GGGGAAGAGG	TGGTGGGAAC	TTCTCAGGAA	AATGGAGAGA	180
AAGAGAACAC	AGACCTGATC	TGAGTAAAC	CACAGGAAAA	CGTACTTCTG	AACAAACCCC	240
ACAGTITTTG	CTTTCAACAA	AGACCCACA	GTCAATGCAG	TCAACATTGG	ATCGATTCTAT	300
ACCATATAAA	GGCTGGAAGC	TTTATTCTC	TGAAGTTTAC	AGCGATAGCT	CTCCTTTGAT	360
TGAGAAGATT	CAAGCATTTG	AAAAATTTT	CACAAGGCAT	ATTGATTTGT	ATGACAGGA	420
TGAATATAGAA	AGAAAGGGAA	GTATTTTGGT	AGATTTTAAA	GAAGTACAG	AAGGTGGTGA	480
AGTAACATAC	TTGATACCG	ATATAGCAAC	TGAAGTAAAG	GATGCACCTG	AGAAAACTTT	540
GGCTTGCATG	GGTTTGGCAA	TACATCAGGT	GTAACTAAG	GACCTTGAAA	GGCATGCAGC	600
TGAGTTACAA	GCCCGAGGAG	GATTGTCTAA	TGATGGAGAA	ACAATGGTAA	ATGTGCCACA	660
TATTATGACA	AGGGGTGACA	ACTATGAGCC	TTTGACACAG	CTCAAGAAATG	TCAGAGCAAA	720
TTACTATGGA	AAATACATTG	CTCTAAGAGG	GACAGTGGTT	CGTGTCACTA	ATATAAGGCC	780
TCTTTGACCC	AAGATGGCTT	TTCTTTGTGC	TGCATGTGGA	GAAATTCAGA	GCTTTCTCTT	840
TCAGATAGGA	AAATACAGTC	TTCCCAACAA	GTGTCTGTG	CCTGTGTGTC	GAGGCAGGTC	900
ATTACTGCT	CTCCGAGCT	CTCCTCTCAC	AGTTACGATG	GACTGGCAGT	CAATCAAAT	960
CCAGGAATTG	ATGTCTGATG	ATCAGAGAGA	AGCAGGTGCG	ATTCCACGAA	CAATAGAATG	1020
TGAGCTTGT	CATGATCTTG	TGGATAGCTG	TGTCCCGGGA	GACACAGTGA	CTATTACTGG	1080
AAATTGCAAA	GCTCTCAATG	CGGAAGAAGG	TTCTCGAAAT	AAGAAATGAC	AGTGTATGTT	1140
CTTTTGTAT	ATTGAAGCAA	ATTCTATTAG	TAATAGCAAA	GGACAGAAAA	CAAGAGGTTT	1200
TGAGGATGGG	TGTAAGCATG	GAATGTTGAT	GGAGTTCTCA	CTTAAGAGACC	TTTATGCCAT	1260
CCAAGAGATT	CAAGCTGAAG	AAAACCTGTT	TAAACTCAIT	GTCAACTCGC	TTTGCCTCTG	1320
CATTTTGGT	CATGAACCTG	TTAAAGCAGG	TTTGGCATT	GCACCTTTTG	GAGGAAGCCA	1380
GAAATACGCA	GATGACAAAA	ACAGAAATCC	AATTCGGGGA	GACCCCAACA	TCCTTGTGTT	1440
TGGAGATCCA	GGCCTAGGAA	AAGTCAAAAT	GCTACAGGCA	GCCTGCAATG	TTGCCCAAGC	1500
TGGCGTGTAT	GTTTGTGGTA	ACACCAGGAC	CACCTCTGGT	CTGACGTTAA	CTCTTTCAAA	1560
AGATAGTTCC	TCTGGAGATT	TTGCTTTGGA	AGCTGGTGCC	CTGGTACTTG	GTGATCAAGG	1620
TATTTGTGGA	ATCGATGAAT	TTGATAAGAT	GGGGAATCAA	CATCAAGCCT	TGTTGGAAAG	1680
CATGGAGCAG	CAAGATTAAT	GTCTTGCTAA	GGCTGGTGTG	GTTTGTAGCC	TTCTCTCAAG	1740
AACCTCCATT	ATTGCTGCTG	CAATCCAGT	TGGAGGACAT	TACAATAAAG	CCAAAAAGT	1800
TTCTGAGAA	TTAAAAATGG	GGAGTGCACT	ACTATCCAGA	TTTATTGTTG	TCTTTATCCT	1860
GTATAGTACT	CCAAATGAGC	ATCATGATCA	CTTACTCTCT	GAACATGTGA	TTGCAATAAG	1920
AGCTGGAAG	CAGAGAACCA	TTAGCAGTGC	CACAGTAGCT	CGTATGAATA	GTCAAGATTG	1980
AAATACTTCC	GTACTTGAAG	TAGTTTCTGA	GAAAGCCATT	TCAGAAAGAC	TAAAGGTGGT	2040
TCTGGAGAA	ACAATAGATC	CCATTCCCA	CCAGCTATTG	AGAAAGTACA	TTGGCTATGC	2100
TOGGCAGTAT	GTGTACCCAA	GGCTATCCAC	AGAAGCTGCT	CGAGTTCTTC	AAGATTTTTA	2160
CCTTGAGCTC	CGGAACAGAA	GCCAGAGGTT	AAATAGCTCA	CCAATCACTA	CCAGGCAGCT	2220
GGAACTCTTG	ATTGCTGTGA	CAGAGGCACG	AGCAAGGTTG	GAATTGAGAG	AGGAAGCAAC	2280
CAAGAGAAC	GCTGAGGATA	TAGTGGAAAT	TATGAAATAT	AGCATGCTAG	GAACCTTACT	2340
TGATGAATTT	GGGAACCTAG	ATTTGAGCG	ATCCAGCAT	GTTTCTGGAA	TGAGCAACAG	2400
GTCAACAGCG	AAAAGATTTA	TTTCTGCTCT	CAACAACGTT	GCTGAAAGAA	CTTATAATAA	2460
TATATTTCAA	TTTCATCAAC	TTGCGCAGAT	TGCCAAAGAA	CTAAACATTG	AGGTTGCTGA	2520
TTTTGAAAT	TTTATGGAT	CACTAAATGA	CCAGGGTTAC	CTCTTGAAAA	AAGGCCCAAA	2580
AGTTTACAG	CTTCAACTTA	TGTAAGAGGA	CTTCAACAG	TTAGGGGCTC	CTGGGTTTAT	2640
TGCAGATTAA	AGCACTCTCA	GTGAAGATAT	GCGTGCAACG	ACAGACAGAG	AGACACACAC	2700
ACACACACAC	ACACACACAC	ACACACACAC	ACACACAGTC	AAATAGTGT	CTCTGAAAAA	2760
TGATGTCCCA	AAGTATTAT	AATAGGAAAA	AAGCATTAAA	TATAATAAAC	TAATTTAAGA	2820

AGTGATAAAG TCTCCAGATG CAGTAGCTCA CACTGTAAATC ACAGTGACTC AGGAGGCTGA 2880
GGTGAGAGGA TTCTTTGAGG CCAGGGTTCCG AGACCAACCT TGGGCAACAT AGCAAGACCC 2940
CATTTCCTAA AAAAAAAAAA AAAAAATTA AACTTAGCTG GGTATGGTGG CACATGCCTA 3000
TAGTCTCAGC TACTTGTGAG GCTGAGGCAG GAGGATTCTT TGAGCCCAGG AGTTTGAGGT 3060
5 TACAGTGAGC CACAATCACA CCAATCACTG CACTCCAGCC TGGGCAATAA AGTAACTCTT 3120
GACTCAAAAA AATAAAAAAA ATTGTAGTGG TAGCCATGTG TTAATTGTGA AATAAAATCT 3180
CCAAAGGGCT AAAAGTAAT TACTTATAAA TTTTATATAG TTGTATTITT GACCTGCCTT 3240
TTATATGTAT GAATATTTC A TAGTTTGA TATCAGATGT AGGCATACAG ACAAATACAT 3300
10 AAACCAATGA ATATATTACA TATTCTGTGT TCCAATAAAA CTTTATTAT GGACACTAAA 3360
ATTTGAATTT CATAAAATTT TCCCATGTCA AGAATACAAA ATACTTGAGT TTTGTTTTTA 3420
GCTATTTAAT AATAGGTCTC ATTATTCCA CAGGCTGTAG TTGTAGTCT TGCTTGAAC 3480
AATAGAAACA GACTGATTAA GCAGGAGAAG TTTTGTGAAA GAATTTTGTG TGGCTCACGG 3540
AATTATTAGA AGGCAGGTGA ACCAGGAGGG TAAGCTTCCA GCAGCAATT GTAAAAACAT 3600
15 GCCTTAGAAT TGGCTAAGG AAGAAGCTGC TGACACTCCA CTGCCACACA GGGCACTGGA 3660
AGAAAGTGCT GCTGCTCTCC TGCCCCACCT TTGCCACTTC TGACAGCAGG ATAGGTAGAA 3720
GAATGCCCCC ACCCGCACCG GAACAGCAAC AAAAGGATTC TGCTAGAGAT GCCTCCCTAA 3780
ATTGCTGAAT TCAAAAGAA AGTTGCATAC AAAGACATCT GATTGAAAAA GGATATGTTA 3840
TATGCCCTT TCAATGGCTG CTAGGGAGTT TTCTGGTTC TACTTTCAGG TGGTGGGATC 3900
20 AATAAGACCA GAATTTCTCA TATGTTGTGA GAGGATTCAA ATGTTACAGG GTTGCCAGCC 3960
AAACTATCAA TACTGTATAA ATCCAACAAA CACTTTGTAA CATAAAGAA CTCAGGAAAT 4020
GTGAACCAAT GTTGGAGAAAT CTACTAAAAT ACGGCTTCCC GCAACCGAAG ATGAATGGAA 4080
AATGTAAATA AAAAGAACTG CCACTGTATA TCAGATGTTT AACTATAGGA CCAGAACTAA 4140
GATGTGAGA CTATTGCCAT AGACCACAAT GTAAATTTT AAGTAGGAA GGAATAATCA 4200
25 GGAATCAAAA GGGGCCAGGT GCAGTGGCTC ACATCTATA TCCAGAGCT TTGGGAGTTC 4260
GAGGAGAGG GATCACTTGA AGCCAGTTT GAGACCAAGC TATGCAACAC ATTGAGACCC 4320
TATCTCTACA AAAAATAGAT TAGCTGGGCA CGGTGGTGCA TGCCTATTGT CCTACCTACT 4380
GTGGAGGCTG AAGTAGGAAA TCACTTGAGC CCGAGAGTT GAGGTTACAG TGAGCTATGA 4440
TTATACCACT GCACTCCAGC CTGGGCAAGA GAGCAAGACC TTGTCTCTT

30 Seq ID NO: 44 Protein sequence:
Protein Accession #: CAB55276.2

1 11 21 31 41 51
MNGEYRGRGF GRGRFQSWKR GRGGNFSKG WREREHRPDL SKTTGKRTSE QTPQFLSTK 60
35 TPQSMQSTLD RPIPYKRWKL YPSEVYSDSS PLIEKIQAPE KFFTRHIDLY DKDEIERKGS 120
ILVDFRELTE GGEVTLNID IATELRDAPE KTLACMGLAI HQVLTQDLER HAAELQAQEG 180
LSNDGETMVN VPHIHARVYN YEPLTQLKNV RANYYGKZIA LRGTVVVRVN IKPLCTKMAP 240
40 LCAACGEIQS FPLPDGKYSL PTKCPVPVCR GRSPFALRSS PLVTMDWQS IKIQELMSDD 300
QREAGRIPT IECELVDHLV DSCVPDGTVT ITGIVKVSNA BEGSRNKNDK CMFLLYLEAN 360
SISNSKGKRT KSSEDGCKHG MLMEFSLKDL YAIQBIQAEZ NLPLKLVNSL CPVIFPGHEL 420
KAGLALALPG GSKYQADDKN RIPIRGDPHI LVVGDPLGLK SQMLQAACVN APRGVVYCGN 480
TTTTSGLTVT LSKDSSSGDF ALEAGALVLG DQICGIDEF DKMGNQHQAL LEAMEQQSIS 540
45 LAKAGVVCSL PARTSIIAAA NPVGGHYNKA KTVSENLKMG SALLSRFDLV FILLDTFNEH 600
HDHLLSEHVI AIRAGKQRTI SSATVARMNS QDSNTSVLEV VSEKPLSERL KVVPGETIDP 660
IPHQLLRKYI GYARQVYPR LSTEAARVLQ DFYLELRKQS QRLNSSPITT RQLESILRLT 720
BARARLELRE EATKEDAEI VEIMKYSMLG TYSDEFNLD FERSQHGSGM SNRSTAKRFI 780
SALNNVAERT YNNIFQFHLQ RQIAKELNIQ VADPENFIGS LNDQGYLLKK GPKVYQLQTM

50 Seq ID NO: 45 DNA sequence
Nucleic Acid Accession #: NM_005416.1
Coding sequence: 149..658

1 11 21 31 41 51
ACCAGATCCC AGAGGCTGAA CACCTCGACC TTCTCTGCAC AGCAGATGAT CCTGAGCAG 60
55 CTGAAGACCA GAAAAGCCAC TAAGACTTTC TGCTTAATTC AGGAGCTTAG AGGATTCTTC 120
AAGAGTGTG TCCACGATCC TTTGAAGCAT GAGTTCTTAC CAGCAGAAGC AGACCTTTAC 180
60 CCCACCACCT CAGCTTCAAC AGCAGCAGGT GAAACAACCC AGCCAGCCTC CACCTCAGGA 240
AATATTGTG CCCACAACA AGGAGCCATG CCACTCAAAG GTTCCACAAC CTGGAACAC 300
AAGATTCCA GAGCCAGGCT GTACCAAGGT CCCTGAGCCA GGCTGTACCA AGGTCCCTGA 360
GCCAGGCTGT ACCAAGGTCC CTGAGCCAGG TTGTACCAAG GTCCCTGAGC CAGGCTGTAC 420
CAAGGTCCT GAGCCAGGTT GTACCAAGGT CCCTGAGCCA GGCTACACCA AGGTCCCTGA 480
65 ACCAGGCAGC ATCAAGGTCC CTGACCAAGG CTTTATCAAG TTCTCTGAGC CAGGTGCCAT 540
CAAGTTCCT GAGCAAGGAT ACACCAAGT TCCTGTGCCA GGCTACACAA AGCTACACGA 600
GCCATGTCTC TCAACGGTCA CTCAGGCCCC AGCTCAGCAG AAGACCAAGC AGAAGTAATT 660
TGTGTGCACAG ACAAGCCCTT GAGAAGCCAA CCACCAAGAT CTGGACACCC TCTTCCCATC 720
70 TGTTCCTGTG TCTTAATGT CTGTAGACCT TGTAAATCAG ACATTGTGAC CCCAAGCCAT 780
AGTCTCTCTC TTAATTGTAT CCTAAAATA CGTACTATAA AGCTTTTGTG CACACACACT 840
CTGAAGAATC CTGTAAGCCC CTGAATTAAG CAGAAAGTCT TCATGGCTTT TCTGCTCTTC 900
GGCTGCTCAG GGTTCATCTG AAGATTCGAA TGAAAAGAAA TGCATGTTTC CTGCTCTTCC 960
CTCATTAAAT TGCTTTAAT TCCA

75 Seq ID NO: 46 Protein sequence:
Protein Accession #: NP_005407.1

1 11 21 31 41 51
MSSYQQKQTF TPPPQLQQQQ VKQPSQPPPQ EIFVPTTKEP CHSKVPQPGN TKIPEPGCTK 60
80 VPEPGCTKVP EPQCTKVPEP GCTKVPEPGC TKVPEPGCTK VPEPGYTKVP EPGSIKVPDQ 120
GFIKFPPEGA IKVPEQGYTK VPVPGYTKLE EPCPSTVTPG PAQQRKQK

85 Seq ID NO: 47 DNA sequence
Nucleic Acid Accession #: Bos sequence

1 11 21 31 41 51
5 GGGTCGTGTG CAGGCGTCCC CGGGCTGTGG ATAATTAGAC ACGTCTCTCC CTCATTGCC 60
AAGGCTCGTT AGAATTGCGC CTAGAGCTGT ATCATGTATT TTCTTTCAAA TTAACCTTGC 120
TTGCAATTAA GCTTAGGGAA CCAGCAACAA AAGCAAACCTT GGCCCGAGGT CGTTCACGCG 180
GAAATGGAT TAGAGAACT TCTTCCCCGA TTTAAGGGGA AAGATTCTGT CGGCCACGCG 240
TTTGGGAAA GTGCCCGAC GCAGAGGGG ACAGCAGGGG AGCAGGAAGC TGCTCACGGT 300
AGTCGGCTT GGGCGCAGG GTGGCCTTCC TCATCTGGGC GATGTGGGCT CCTAGAAGAG 360
TAAGGATAAC ATCTCGAAA TGACTTCTGT ACGGTTTGAG CCAACTGCA CACTCATGAC 420
10 TTGAGCTGC CCTGTGAGT TACAGTTTAC CAAACACATT CATGAACATA ATCTCAITTA 480
CTAAAAACTT TGTGAGAAAT TTCTTTTACT AAAATTTTCT CTTATTACAA A

Seq ID NO: 48 DNA sequence:
Nucleic Acid Accession #: CAT cluster

15 1 11 21 31 41 51
20 TTCCAAATTT TTTTTTTTGT AATAAGAAAA AATTTTAGTA AAAGAAAATT CTCACAAAGT 60
TTTATGTAAA TGAGATTATG TTCATGAATG TGTITGGTAA ACTGTAACTC CACAGGGCAG 120
CTCCAAGTCA TGAGTGTGCA GTTGGGCTCA AACCGTACAG AAGTCATTTC CAGGATGTTA 180
TCCTTACTCT TCTCGAGGCC CACATCGCCC AGATGAGGAA GGCCACCGCT GCCGCCAAGC 240
CCGACTACCG TGAGCAGCTT CCTGCTCCCC TGTGCTCGCC TCTGCGGTGG GGGCACTTTC 300
25 CCCAAGGCGC TGGCCGAGG AATCTTCCCC CTTAAATCGG GGAAGAAAGT TCTCTAATCC 360
ATTTTCGCGG TGAACGACCT CGGGCCAAAG TTGCTTTTGT TGCTGGTTCC CTAAGCTTAA 420
TTGCAAGCAA AGTTAATTTG AAAGAAAATA CATGATACAG CTCTAGGGCG AATTCTAACG 480
AGCCTTGGGC AATGAGGGAA GAACGTGTCT AGTTATCCAC AGCCCGGGGA CGCCTGCACA 540
CGACGCT

30 Seq ID NO: 49 DNA sequence
Nucleic Acid Accession #: CAT cluster

35 1 11 21 31 41 51
40 TCTTCTCTCT GCTGCTCGTT TGTCTCTCCT GTGCTCTTCT TCTTCTTTTC CCTCGCGGCT 60
CCTGCGCAGC TCTGTTGTCT CTCTCTGTAT GGGCGGGGGC GGGAGAGGCT GACCGGTGAG 120
ACCGTAGACC CGAAACCATT GGGTGTGACA AGCGGTGCGC CGGCTTTTTT GGGAGAACCC 180
GACACATGCA GACCAGTTTT CCTGGAACNG CATGACCATG TTAITACTAT GGGCCGCCCTC 240
45 CCCAACCAA GTGTTTAAAA CTTTTTAGGG CACCCCCAAA ATTTTTTTTT TTTTTTTTTT 300
TTCAATTTAA AAACCTTAAT ATTTATATTA AATACAAAGA TACCCAAACC CTTTATGCTT 360
CTTTCCTGTA TCTGTGTCTT TTTTCTTTGA CAGCATCTCC ATTTTTTTTC TGCTGCTTCA 420
TCGCTGTAGC CATGGAAATC GCTTTCATTA TTATGGTAGC AATATGGAGT GCTGTATTCC 480
TAAAGAAACT GACACAGGAG AATCACTTGA ACTTGGGAGG CAGAGTTTGC AGTGAGCCGA 540
GATTGAACCA GTGCACCTCCA GCCTTGGCAG CGGAGCAAGA TTCTGTGACA GTTCTCTGAG 600
45 TGCTGGTATC GTCTCTGAGC CCATCTCTCG GTTCCATGTC GCTGCCAGGG AGGGTCTGG 660
GAGTGGGGA GAGCTGTCTT ATATATCCGG GTGAAGCTCA GCTGTGGCAC ACCTTGGATG 720
CCGGTCTCT CCTGCCCCCG GGGACCTAGT ATTTTGGCCA CGAGTGTACA CCAACAAAG 780
GAGACAGCAT CATTTATGAG CCTGCAGCAT CCACGCTACT GCTGTATCCA GTTTCATTG 840
50 ACTG

Seq ID NO: 50 DNA sequence
Nucleic Acid Accession #: L05187
Coding sequence: 1991..2260

55 1 11 21 31 41 51
60 CTGCAGGGAG GCAGGTAGAA AAGGCTTTTG GGTTTTCAGG TGGGGGGCAG TCTAGCCTGA 60
TCAGAAAGGA GGAAAGGGCC AGGGCAGATG TCTGGGTGGA GTGAAGGGAA AAAGTGATCC 120
CAGAAGAAAG ATTAGCCCCT GAAAGTCCCT GAAGTAGGAG AAGGGTAAAG GTGTGGTTGG 180
65 TGAGAGGAAG CAGGTTTTC CAGATTAGCA ACCAGTCAGG GGGAGGAAGG TGAGAGTGGG 240
AGAGTCATAA GTAAATTAAT CTGAATGTGT GTAGTTAAT GGAATTGGGA AAAAGATGGG 300
GGAAATGGAT GGAAGGTCCT GGACTCTGAG ACAAGGGGTC TATAATCAGT CCAITTCATT 360
ATTTCTAGCT TCCACCTTCA CCAAGGCAGA CAAGGAGGSC CCACTCAGC TCTCTGCTC 420
CCCCCTCCCT TCCCACTTAT TCATGTGTGC AAGAGTGGCC TGTCCACAG AACACGGGGA 480
65 ACAACCATCT CAATGACAAG GACAGCAGGT GGCAGGCTC AACAGGACTC AGATGTCCCC 540
CCAGGGTTAA CTCATGAAC CCTCCATGAA GCCTGCTGCT CACCCCTCCC TCAAGGCAAG 600
CCCTGCACCT GGGTCTGAGG ATGAGGGTGG CAGTGAAAT TAGGCCAGTG ACATCATTTT 660
CAGCCAGCTA GTGCCAAAAA ATATCAGGTG GTGTTTCATCA AATAAGCCGA GCCAACCGGT 720
70 GATGAGGATG GTAGTGTGAG TCATGTGTGA CAGGTGAGGA ATGAAAAAG AGTGCCCGAG 780
AGCTTCTATT TCCTTGAGGC AGGGCTCATT CATCTATAA AAGCCAGCTG GCCATTGCCT 840
TCACACCAA CCCAAGGGAC CACACAGCCC ATTCTGCTCC GTATACCAGG TAAGTCTCTG 900
ATTGCAACAA ACTGGCAATT CTAGTGTACT TTTTCATTAT TAGAAATTAG CTAAGGCCAA 960
ATATGTGTAA CAGAGTTAAT CCAGGGTTTC AATGGGAGAT AGAGAATAGT GGAATATCTT 1020
75 TATTTTAAGT TAAATTACAG TCTGGATTG AAAGGACCTT AGAGATGGTT AGGGCTCCCC 1080
CCTCAGTAGA TAGTCATTGA ACTGGGATC CTGGAGAAGA TTGTTCAAAT GCCCATGGGA 1140
AGTTCATAGC AGAAGTAGAA CTCAGGCCAG AGCACTCTCA GTAACACTGC AATTTCCCCC 1200
TGACAAGATA TTTATAGAAA TTTTAATTAT TTAGATGGAT CTCTACTGAG CATTTATTCC 1260
ATTTAAGGCA GTATGCTTAG CACTTTGGAC AAATCAATGC CCTAACGTAC TTACTTAACA 1320
AACATAAATC CTAGCAGGAA GGTAAATCAT ATATATAAAT AAATGAAATG CAAAGTAGAT 1380
80 AGTAATTGGC ATGACGGAGA TGGGCAGAGA AGGGCTGTGC ACTTTTGGGA GACTTGCTCA 1440
AGGAGACCTC TAGGGTGTCA AGTGATGTGA GCTATGATGG AGGGGTATTT GGACAAGCAG 1500
AGATGGGAAG AAAAGCATTT GGAAGGGACT GTGTAAAGCAC AGACCAGAAG CAAACACATA 1560
GAGGCTTAGA TGAATTATAA GCCATCTTAT AAGTCACAGG CTTTCTACAT GGTACTAGGA 1620
GAGGAAGTGT GTCTGATGCC ATTTTCCAAA AGACCTAATA TGCGGACCTC ATGTCCCTCA 1680
85 GAAGCCAGCT TTAGTAGGGC ATTTTCCAG AACAGATATA AGGTGCCCTG GGTAGGAAGG 1740
GAGCCAGAA GAGAACTCCA ATAAAATGGA GCAGAAAGAA TTGCCTTTTA GCTCTCTCTC 1800
TTCAAGGGC CTGAAAATTA TCCAAGCTTA TTTCAATTTT AAATGTAATG GGGGAGCTAA 1860

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GGGAGATGAA AGGCTTCTCT TTCTAAAGGG TCCTGAAATA AAATCTGTTT GGCATTGAAT 1920
TTGTATCCAT CTTTCTTTAA TTGAATCACT GTGTCAAGTT TCTGTCTCTA GAAAAAACA 1980
CATTTGAAGC ATGAATTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA 2040
GCAGCAGCAG GTGAACAAC CTGGCCAGCC TCCACCCAG GAAACATGCA TCCCAAAAC 2100
CAAGGAGCCC TGCCAACCA AGGTGCCCTGA GCCCTGCCAC CCCAAAGTGC CTGAGCCCTG 2160
CCAGCCCAAG ATTCAGAGC CCTGCCAGCC CAAGGTGCCT GAGCCCTGCC CTTCAACGGT 2220
CACTCCAGCA CCAGCCCAGC AGAAGACCAA GCAGAAGTAA TGTGTTCAC AGCCATGCC 2280
TTGAGGAGCT GGCCACTGGA TACTGAACAC CCTACTCCAT TCTGCTTATG AATCCCATTT 2340
GCCTATTGAC CTTGCACTTA GCATGCTGTC ACCCTGAATC ATAATCGCTC CTTTGCACCT 2400
CTAAAAAGAT GTCCCTTACC CTCATTCTGG AGGCTCCTGA GCCTCTGCGT AAGGCTGAAC 2460
GTCTCACTGA CTGAGCTAGT CTTCTTGTG CTGGGTGCA TTTGAGGATG GATTTGGGGA 2520
AGGTCAGTG ACCATCCCTA G

Seq ID NO: 51 Protein sequence:
Protein Accession #: AAC26838

1 11 21 31 41 51
MNSQQQKQPC TFPFPQPPQQ VKQPCQPPPO EPCIPKTKEP CQPKVPEPCN PKVPEPCQPK 60
IPEPCQPKVP EPCPSTVTPA PAQQKTKQK

Seq ID NO: 52 DNA sequence
Nucleic Acid Accession #: NM_002638.1
Coding sequence: 120-473

1 11 21 31 41 51
CAATACAGCT AAGGAATTAT CCCTTGTAAG TACCACAGAC CGCCCTGGA GCCAGGCCAA 60
GCTGGACTCG ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGCTGGGAGC CTGGTTCTAG 180
AGGCAGCTGT CACGGAGATT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240
TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300
CGCAAGAGCC AGTCAAAAGT CCAGTCTCCA CTAAGCCTGG CTCTGCCCC ATTATCTTGA 360
TCGGTGGCC CATGTTGAAT CCCCTAACC GCTGCTTGAA AGATACTGAC TGCCAGGAA 420
TCAAGAAAGT CTGTGAAGGC TCTTGCGGGA TGGCCTGTT CGTTCGCCAG TGAAGGGAGC 480
CGGTCTTGC TGCACCTGTG CGGTCCCGAG AGCTACAGGC CCCATCTGCT CTAAGTCCC 540
TGCTGCCCTT CCCCTTCCCA CACTGTCCAT TCTTCTCC ATTCAAGATG CCCACGGCTG 600
GAGCTGCTC TCTCATCCAC TTTCATTA A

Seq ID NO: 53 Protein sequence:
Protein Accession #: NP_002629.1

1 11 21 31 41 51
MRASSPLIVV VFLIAGTLVL EAAVTGVPEVK GQDTVKGKVP FNGQDPVKGO VSVKGQDKVK 60
AQEPVKGFVS TKPFGSCPIL IRCAMLNPPN RCLKDTDCPG IKKCEBSCG MACFVPQ

Seq ID NO: 54 DNA sequence
Nucleic Acid Accession #: NM_019618
Coding sequence: 75-584

1 11 21 31 41 51
GGCAGAGCC ACGATTCACT CCCCTGGACT GTAGATAAAG ACCCTTCTT GCCAGGTGCT 60
GAGACAACCA CACTATGAGA GGCCTCCAG GAGAGCTGA TGGTGGAGGA AGGGCGTCT 120
ATCAATCAAT GTGTAAACCT ATTACTGGGA CTATTAATGA TTTGAATCAG CAAGTGTGGA 180
CCCTTCAGGG TCAGAACTTT GTGGCAGTTC CACGAAGTGA CAGTGTGACC CCAGTCACTG 240
TTGCTGTTAT CACATGCAAG TATCCAGAGG CTCTTGAGCA AGGCAGAGGG GATCCCATTT 300
ATTTGGGAAT CCAGAATCCA GAAATGTGTT TGTATTGTGA GAAGGTGGA GAACAGCCCA 360
CATTCAGCT AAAAGAGCAG AAGATCATGG ATCTGTATGG CCAACCCGAG CCCGTGAAAC 420
CCTTCCTTTT CTACCGTGCC AAGACTGGTA GGACCTCCAC CCTTGAGTCT GTGGCCTTCC 480
CGGACTGGTT CATTCGCTCC TCCAAGAGAG ACCAGCCCAT CATTCTGACT TCAGAACTTG 540
GGAAGTCATA CAACACTGCC TTTGAATTAA ATATAATGA CTGAACCTAG CCTAGAGGTG 600
GCAGCTGGT CTTGTCTTA AAGTTCTGG TTCCAATGT GTTTTGTCT ACATTTTCTT 660
AGTGTCAATT TCACCGTGGT GCTGAGACAG GGGCAAGCCT GCTGTATCA TCTCATTTTA 720
TAATGAAGAA GAAGCAATTA CTTATAGCA ACTGAAGAAC AGGATGTGGC CTCAGAAGCA 780
GGAGAGCTGG GTGTATAAG GCTGTCTCT CAAGCTGGTG CTGTGTAGGC CACAAGGCAT 840
CTGCATGAGT GACTTTAAGA CTCAAAGACC AAACACTGAG CTTTCTCTA GGGGTGGGTA 900
TGAAGATGCT TCAGAGCTCA TGCGCGTTAC CCAAGATGGC ATGACTAGCA CAGAGCTGAT 960
CTCTGTTTCT GTTTTGCTTT ATTCCCTCTT GGGATGATAT CATCCAGTCT TTATATGTTG 1020
CCAATATACC TCATTGTGTG TAATAGAACC TTCTTAGCAT TAAGACCTTG TAAACAAAAA 1080
TAATCTTGT GTTAAGTTAA ATCAATTTTG TCCTAATTGT AATGTGTAAT CTTAAAGTTA 1140
AATAAACTTT GTGATTATTAT ATAATAAAAA AAAAAAAAAA AAA

Seq ID NO: 55 Protein sequence:
Protein Accession #: NP_062564

1 11 21 31 41 51
MRGTPGDADG GSRVAVYQSMC KPITGTINDL NQQVWTLQGG NLVAVPRSDS VTPVTVAVIT 60
CKYPEALEQG RGDPIYLGIIQ NPENCLYCEK VGEQPTLQLK EQKIMDLVNG PEPVKEPFLFY 120
RAKTRSTSL ESVAFPDWFI ASSKRDQPII LTSELGKSYN TAFELATIND

Seq ID NO: 56 DNA sequence
Nucleic Acid Accession #: NM_003125
Coding sequence: 65-334

```

1      11      21      31      41      51
|      |      |      |      |      |
5  AGCAGTTCTA AGGGACCATTA CAGAGTATTC CTCTCTTCAC ACCAGGACCA GCCACTGTG 60
   CAGCATGAGT TCCCAGCAGC AGAAGCAGCC CTGCATCCCA CCCCCTCAGC TTCAGCAGCA 120
   GCAGGTGAAA CAGCCTTGCC AGCCTCCACC TCAGGAACCA TGCAATCCCA AAACCAAGGA 180
   GCCCTGCCAC CCCAAGGTGC CTGAGCCCTG CCAACCCAAA GTGCGTGAGC CCTGCCAGCC 240
   CAAGCTTCCA GAGCATGACC ACCCAAGGT GCCTGAGCCC TGCCCTTCAA TAGTCACTCC 300
   AGCACCAACC CAGCAGAAGA CCAAGCAGAA GTAATGTGGT CCACAGCCAT GCCCTTGAGG 360
10  AGCGGGCCAC CAGATGCTGA ATCCCCATAT CCATTCTGTG TATGAGTCCC ATTTGCCTTG 420
   CAATTAGCAT TCTGTCTCCC CCAAAAAGA ATGTGCTATG AAGCTTTCTT TCCTACACAC 480
   TCTGAGTCTC TGAATGAAGC TGAAGTCTT AGTACCAGAG CTAGTTTTC A GCTGCTCAGA 540
   ATTCACTCGA AGAGAGACTT AAGATGAAAG CAAATGATTC AGCTCCCTTA TACCCCATTT 600
15  AAATTCATT TCAATTCCA

```

Seq ID NO: 57 Protein sequence:
Protein Accession #: NP_003116

```

20      1      11      21      31      41      51
|      |      |      |      |      |
   MSSQQQKQPC IPPPQLQQQQ VKQPCQPPFQ EPCIPKTKEP CHPKVEPECH PKVPEPCQPK 60
25  LPEPCHPKVP EPCPSIVTPA PAQKTKQK

```

Seq ID NO: 58 DNA sequence
Nucleic Acid Accession #: NM_001793.2
Coding sequence: 71-2560

```

30      1      11      21      31      41      51
|      |      |      |      |      |
   AAAGGGGCAA GAGCTGAGCG GAACACCGGC CCGCGGTGCG GGCAGCTGCT TCACCCCTCT 60
   CTCTGCAGCC ATGGGGCTCC CTCGTGGACC TCTCGGCTCT CTCCTCCTTC TCCAGGTTTG 120
   CTGGCTGCAG TGCGCGGCGT CCGAGCCGCT CCGGGCGGTC TTCAGGGAGG CTGAAGTGAC 180
35  CTGGAGGCGG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAG TATTCTATGG 240
   CTGCCCTGGG CAAGAGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300
   TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360
   ATCCAAAGCT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420
   TGAAAATGGC AAGGGTCCCT TCCCCCAGAG ACTGAATCAG CTCAGTCTTA ATAAAGATAG 480
40  AGACACCAAG ATTTTCTACA GCATCACGGG GCGGGGGGCA GACAGCCCCC CTGAGGGTGT 540
   CTTCGCTGTA GAGAAGGAGA CAGGCTGGTT GTTGTGAAT AAGCCACTGG ACCGGGAGGA 600
   GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTGAGAG AATGGTGCTC CAGTGGAGGA 660
   CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720
   GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCCTACCA GGTACTTCTG TGATGCAGGT 780
45  GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAAT GGGGTGGTGG CTTACTCCAT 840
   CCATAGCCAA GAACCAAAGG ACCCACAGA CCTCATGTTC ACCATTACCC GGAGCACAGG 900
   CACCATCAGC GTCACTCTCA GTGGCTTGA CCGGGAAAAA GTCCCTGAGT ACACACTGAC 960
   CATCCAGGCC ACAGACATGG ATGGGGACGG CTCACCAACC ACGGCAGTGG CAGTAGTGGA 1020
   GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080
50  GCCTGAGAAAT GCAGTGGGCG ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140
   CAACCTACCA GCGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACG GGGACCAATT 1200
   TACCATCACC ACCACCCCTG AGAGCAACCA GGGCATCTCT ACAACAGGA AGGGTTTGA 1260
   TTTTGAGGCC AAAAACCAGC ACACCCCTGTA CGTTGAAGTG ACCAAGAGG CCCCTTTTGT 1320
   GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAC GTGGAGGATG TGAATGAGGC 1380
55  AOCCTGTGTT GTCCCAACCT CCAAAGTCGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440
   GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAGA TCAGTACCG 1500
   CATCTTGAGA GACCCAGCAG GGTGGCTAGC CATGGACCCA GACAGTGGGC AGGTCAACG 1560
   TGTGGGCACC CTCGACCGTG AGGATGAGCA GTTTGTGAGG AACCAATCT ATGAAGTCAT 1620
   GTCTTGTGCC ATGGCAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGTCTAC 1680
60  ACTGATTGAT GTCAATGACC ATGGCCAGT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740
   CCAAAGCCCT GTGCGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800
   CCCCTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860
   GGAAGGTGAC ACATGGTCT TGTCCCTGAA GAAGTTCCCT AAGCAGGATA CATATGACGT 1920
   GCACCTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980
65  GTGOGACTGC CATGGCCATG TCGAAACCTG CCCTGGAACC TGGAAAGGAG GTTTCATCTC 2040
   CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCTCTCTG CTGGTCTGCT TTTTGTGTTG 2100
   GAGAAAGAAG CGGAAGATCA AGGAGCCCTC CTTACTCCCA GAAGATGACA CCGTGACAA 2160
   CGTCTTCTAC TATGGCGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220
   GCTCCACCGA GGTCTGGAGG CAGGGCCGGA GGTGGTTCTC CGCAATGAOG TGGCACCAAC 2280
70  CATCATCCCG ACACCCATGT ACGTCTCTG GCCAGCCAAC CCAGATGAAA TCGGCAACTT 2340
   TATAATTGAG AACCTGAAGG CGGCTAACAC AGACCCCAAC GCCCGCCCTC ACGACACCT 2400
   CTGTGTGTTT GACTATGAGG GCAGCGGCTC CGACGCGCGT TCCCTGAGCT CCCTCACCTC 2460
   CTCCGCTCTC GACCAAGACC AAGATTAAGA TTATCTGAAC GAGTGGGGCA GCGCTTCAA 2520
   GAAGCTGGCA GACATGTACG GTGGCGGGGA GGACGACTAG GCGGCTGCC TGCAGGCTG 2580
75  GGGACCAAAC GTCAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCTC TTCAGCTGAG 2640
   GACTTCGAGG CTTGTGAGG AGTGGCGTGA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700
   ACGTTAGAGT GGTGTCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTTGACTTC 2760
   AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCCCTCAG AGGCCAAGTT TCCAGAGCC 2820
   TCTTACCTGC CGTAAATATG TCAACCTGT GTCTTGGGCC TGGGCTGCT GTGACTGACC 2880
80  TACAGTGGAC TTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTT 2940
   TTTTITTAAT GCTATCTTCA AAACGTTAGA GAAAGTCTT CAAAAGTGCA GCCCAGAGCT 3000
   GCTGGGCCCA CTGGCGTCC TGCAATTCTG GTTTCAGAC CCCAATGCCT CCCATTGGA 3060
   TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGGTTGCC CCTTATTTT TATTTTCCCT 3120
85  GTTGGTGGC TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTTTAT 3180
   TAAAGAAACT TTTCACGAA AAAAA

```

Seq ID NO: 59 Protein sequence:

Protein Accession #: NP_001784.2

1 11 21 31 41 51
5 MGLPRGFLAS LLLLVQVCWLQ CAASEPCRAV FREAEVTLA GGAQEPGQA LGKVFMGCPG 60
QEPALFSDTN DDPFVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVFENG 120
KGFPPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLEN KPLDREEIAK 180
YELFGHAYSE NGASVEDPMN ISIIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240
DEDDAIYTYN GVVAYSISHSQ EPKDPHDLMP TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300
10 TDMDGDGSTT TAVAVVEILD ANDNAPMDFP QKYEAVPEN AVGHEVQRLT VTDLDAPNSP 360
AMRATYILMG GDDGDHFTIT THPESNQGLL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420
PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEFVC VYTAEDPDKE NQKISYRILR 480
DPAGWLAMPD DSGQVAVGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLLTLD 540
VNDHGFVPEP RQITICNQSP VRQVLNITDK DLSPTSPFQ AQLTDDSDIY WTAEVNEEGD 600
15 TVVLSLKKFL QDQTYDVHLS LSDHGNKEQL TVIRATVCDG HGHVETCPGP WKGGFILPVL 660
GAVLALLPLL LVLLLLVRKK RKIKEPLLL EDDTRDNVFP YGEEGGGEED QDQYDITQLHR 720
GLEARPEVVL RNDVAPTIP TPMYRPRPAN PDEIGNFIE NLKAANTDPT APPYDTLLVF 780
DYGSGSDAA SLSSLTSSAS DQDQDYDYLW EWGSRPKKLA DMYGGGEDD

Seq ID NO: 60 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 162-428

1 11 21 31 41 51
25 GCGTCCGTT GCGCGCGGAT TCGAACGTT GGAAGTGGT TTTCTGCCT GAAGAAGCGT 60
CATACGGACC GGATTTGTTT CGCTGGCCCA GTGTCCCGG AGCTTGTGTG CGATACAGAG 120
AGCACTCGG AAGCTGAGGC AGCTGGTACT TGACAGAGAG GATGGCGCTG TCGACCATAG 180
TCTCCAGAG GAAGCAGATA AAGCGGAAG CTCCCGTGG CTTTCTAAG CGAGTCTTCA 240
30 AGCGAAAGAA GCGTCAACTT CGTCTGGAGA AAAGTGGTGA CTTATTGGTC CATCTGAAC 300
GTTTACTGTT TGTTCATCGA TTAGCAGAG AGTCCAGGAC AAACGCTTGT GCGAGTAAAT 360
GTAGAGTCAT TAACAGGAG CATGTACTGG CCGCAGCAAA GGTAATCTA AAGAAGAGCA 420
GAGGTTAGAA GTCAAGAAG ATATTCTTGA AAGTTATGAT GCATTCTTTT GGGTGGTAA 480
AGATCATAAA GACATTTTTT ACACATCAGT TAATATGGGA TTATTAAATA TTGG

Seq ID NO: 61 Protein sequence:

Protein Accession #: Eos sequence

1 11 21 31 41 51
40 MALSTIVSQR KQIKRKAPRG FLKRVFKRK POLRLEKSGD LLVHLNCLLF VHLAEESRT 60
NACASKRCVI NKEHVLAAAK VILKRSRG

Seq ID NO: 62 DNA sequence

Nucleic Acid Accession #: NM_000094.2

Coding sequence: 99-8933

1 11 21 31 41 51
50 GGGCTGGAGG GCGCTGGGCG TCGGACCTGC CAAGGCCACC GCAGGGGGGA GCAAGGGACA 60
GAGGCGGGGG TCCTAGCTGA CGGCTTTTAC TGCCTAGGAT GACGCTGCGG CTTCTGGTGG 120
CCGCGCTCTG CCGCGGATC CTGGCAGAGG CGCCCGAGT GCGAGCCAG CACAGGGAGA 180
GAGTGACCTG CAGCGCGCTT TACGCGCTG ACATTGTGTT CTTACTGGAT GGCTCCTCAT 240
CCATTGGCCG CAGCAATTTT CCGAGGTGCC GCAGCTTTCT CGAAGGGCTG GTGCTGCCTT 300
55 TCTCTGGAGC AGCCATGTGA CAGGGTGTGC GCTTTGCCAC AGTCAGTAC AGCGATGACC 360
CAGGACAGA GTTCGGCTG GATGCACTTG GCTCTGGGGT TGATGTGATC CGGCCCATCC 420
GTGAGCTTAG CTACAAGGGG GGCACACTC GCACAGGGGC TGCAATTCTC CATGTGGCTG 480
ACCATGTCTT CCTGCCCCAG CTGGGCGGAC CTGGTGTCCC CAAGGTCTGC ATCCTGATCA 540
CAGACGGGAA CTCCAGGAC CTGGTGGACA CAGCTGCCCA AAGGCTGAAG GGGCAGGGGG 600
60 TCAAGCTATT TGCTGTGGGG ATCAAGAATG CTGACCCCTG GAGAGCTGAAG CGAGTTGCCT 660
CAGACCAAC CTCCGACTTC TTCTTCTTGG TCAATGACTT CAGCATCTTG AGGACACTAC 720
TGCCCCCTGT TTCCCGGAGA GTGTGCACGA CTGCTGGTGG CGTGCCCTGT ACCCGACCTC 780
CGGATGACTC GACCTCTGCT CCAAGAGACC TGSTGCTGTC TGAGCCCAAG AGCCAATCCT 840
TGAGAGTACA GTGGACAGCG GCCAGTGGCC CTGTGACTGG CTACAAGGTC CAGTACACTC 900
65 CTCTGACGGG GCTGGGACAG CCACTGCCGA GTGAGGGGCA GGAGGTGAAC GTCCAGCTG 960
GTGAGACCAG TGTGGGCTG CCGGGTCTCC GGCCACTGAC CGAGTACCAA GTGACTGTGA 1020
TTGCCCTCTA CGCCACAGC ATCGGGGAGG CTGTGAGCGG GACAGCTCGG ACCACTGCCC 1080
TAGAAGGGCC GGAAGTGAAC ATCCAGAATA CCACAGCCCA CAGCCTCTG GTGGCCTGGC 1140
GGAGTGTGCC AGGTGCCACT GGCTACCGTG TGACATGGCG GGTCTCAGT GGTGGGCCCC 1200
70 CACAGCAGCA GAGCTGGGC CTTGGGCAGG GTTCAAGTGT GCTGCGTGAC TTGGAGCCTG 1260
GCACGAGTGA TGAGGTGACC GTGAGCAACC TATTGGCCG CAGTGTGGGG CCGGCCACTT 1320
CCCTGATGCG TCGACTGAC GCTTCTGTTG AGCAGACCTT CGCGCCGGTC ATCCTGGGCC 1380
CCACATCCAT CCTCTTTTCC TGGAACTTGG TGCCGTGAGG CCGTGGCTAC CGGTTGGAAT 1440
75 GCGCGGCTGA GACTGGCTTG GAGCCACCGG AGAAGGTGGT ACTGCCCTCT GATGTGACCC 1500
GCTACCAATT GGTGGGCTG CAGCCGGGCA CTGAGTACCG OCTCACACTC TACACTCTGC 1560
TGGAGGGCCA CGAGGTGGCC ACCCCTGCAA CGTGGTTC CACTGGACCA GAGTGTCTG 1620
TGAGCCCTGT AACAGACCTG CAAGCCACCG AGCTGCCCGG GCAGCGGGTG CGAGTGTCTT 1680
GGAGCCCACT CCTGTGTGCC ACCCAGTACC GCATCATTGT GCGCAGCACC CAGGGGGTGG 1740
AGCGGACCTT GGTGCTTCTT GGGAGTCAGA CAGCAATCGA CTTGGATGAC GTTCAGGCTG 1800
80 GGCTTAGCTA CAGTGTGGG GTGTCTGCTC GAGTGGGTCC CCGTGAAGGG AGTGCCAGTG 1860
TCCTCACTGT CCGCGGGGAG CCGGAAACTC CACTTGTCTG TCCAGGGCTG CCGGTTGTGG 1920
TGTCAAGTGC AAGCGAGTGT AGGTGTGGCT GGGGACCGGT CCTGGAGGCG AGTGATTTC 1980
GGATTAGCTG GAGCAGAGGC AGTGGTCCGG AGTCCAGCCA GACACTGCCC CCAGACTCTA 2040
CTGCCACAGA CATCAGAGG CTGCAGCTG GAACCACTTA CCAGGTGGCT GTGTCGGTAC 2100
85 TGGAGGCTAG AGAGAGGGG CCTGCTGAG TCATCGTGGC TCGAAAGGCA CCACTGGGCC 2160
CAGTGAGGAC GGTCCATGTG ACTCAGGCCA GCAGCTCATC TGTCAACATT ACCTGGACCA 2220
GGTTCCTG CGCCACAGGA TACAGGGTTT CTTGGCACTC AGCCACGGC CCAGAGAAAT 2280

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

CCCAGTTGGT TTCTGGGGAG GCCACGGTGG CTGAGCTGGA TGGACTGGAG CCAGATACTG 2340
 AGTATACGGT GCATGTGAGG GCCCATGTGG CTGGCGTGA TGGGCCCCCT GCCTCTGTGG 2400
 TTGTGAGGAC TGCCCTCTGAG CCTGTGGGTC GTGTGTGCGAG GCTGCAGATC CTCAATGCTT 2460
 CCAGCGAGCT TCTACGGATC ACCTGGGTAG GGGTCACTGG AGCCACAGCT TACAGACTGG 2520
 CCTGGGGCCG GAGTGAAGGC GGCCCATGA GGACCCAGAT ACTCCAGGA AACACAGACT 2580
 CTGACAGAGT CCGGGGTCTC GAAGGTGGAG TCAGCTACTC AGTGCAGTGG ACTGCACCTG 2640
 TCGGGGACCG CGAGGGGACA CCTGTCTCCA TTGTGTGTAC TACGCGGCTC GAGGCTCCGC 2700
 GAGCCCTGGG GAGGCTTAC GTGGTGAGC GCGGGGAGCA CTGCTGAGG CTGCGCTGGG 2760
 AGCCGCTGCC CAGAGCGCAG GGCTTCCTTC TGCACTGSCA ACCTGAGGGT GGCCAGGAAC 2820
 AGTCCCGGGT CCTGGGGCCC GAGCTCAGCA GCTATCACTT GGAAGGGCTG GAGCCAGCGA 2880
 CACAGTACCG CGTAGGCTG AGTGTCTTAG GGCCTGGCTG AGAAGGGCCC TCTGCAGAGG 2940
 TGACTGCGCG CACTGAGTCA CCTCGTGTTC CAAGCATTGA ACTACGTGTG GTGGACACCT 3000
 CGATCGACTC GGTGACTTTG GCCTGGACTC CAGTGTCCAG GGCATCCAGC TACATCTTAT 3060
 CCTGGCGGCC ACTGAGAGGC CTGGGCCAGG AAGTGCTGG GTCCCGCAG ACACCTCCAG 3120
 GGATCTCAAG CTCCAGCGG GTGACAGGGC TAGAGCTTGG CGTCTCTTAC ATCTTCTCCC 3180
 TGACGCTGCT CCTGATGGT GTGCGGGGTC CTGAGGCATC TGTCAACAG ACGCCAGTGT 3240
 GCCCCGCTGG CCTGGCGGAT GTGGTGTTC TACCACATGC CACTCAAGAC AATGCTCACC 3300
 GTGCGGAGGC TAGGAGGAGG GTCTGGAGG GTCTGGTGTG GGCACCTGGG CCTCTTGGGC 3360
 CACAGGCACT TCAGGTGGG CTGCTGTCTT ACAGTCACTG GCCCTCCCA CTGTCTCCAC 3420
 TGAATGGCTC CCATGACCTT GGCATTATCT TGCAAAGGAT CCGTGACATG CCCTACATGG 3480
 ACCAAGTGG GAACAACCTG GGCACAGCCG TGGTCAACAG TCACAGATAC ATGTGGGAC 3540
 CAGATGCTCC TGGCGCGCCG CAGCAAGTAC CAGGGGTGAT GGTCTGCTA GTGGATGAAC 3600
 CCTTGAGAGG TGACATATTC AGCCCATCC GTGAGGCCCA GGTCTCTGGG CTAAATGTGG 3660
 TGATGTTGGG AATGGCTGGA GCGGACCCAG AGCAGCTGCG TCGCTTGGCG CCGGTATGG 3720
 ACTCTGTCCA GACCTTCTTC GCGTGGATG ATGGGCCAAG CCTGGACCAG GCAGTCAGTG 3780
 GTCTGGCCAC AGCCCTGTGT CAGGCATCTC TCACTACTCA GCCCGGGCCA GAGCCCTGCC 3840
 CAGTGTATTG TCCAAAGGGG AACTTGGAGA GATGGGCTG AGAGGACAAG 3900
 TTGGGCTTCC TGGCAGCCCT GGCCTCCCGG GCAGGACCGG TGCTCCCGGC CCCCAGGGGG 3960
 CCCCTGGAAG TGCCACTGCC AAGGGCGAGA GGGGCTTCCC TGGAGCAGAT GGGCTTCCAG 4020
 GCAGCCCTGG CCGCGCGGGG AATCCTGGGA CCCCTGGAGC CCCTGGCCTA AAGGGCTCTC 4080
 CAGGGTGGCC TGCCCTCTGT GGGGACCCCG GAGAGCGAGG ACCTCGAGGC CCAAAGGGGG 4140
 AGCCGGGGGC TCCCGGACAA GTCATCGGAG GTGAAGGACC TGGGCTTCTT GGGCGGAAG 4200
 GGGACCCCTG ACCATCGGCG CCCCTGGAC CTCTGTGACC ACTGGGGGAC CCAGGACCCC 4260
 GTGGCCCCCC AGGGCTTCTT GGAACAGCCA TGAAGGGTGA CAAAGGGGAT CGTGGGGAGC 4320
 GGGGTCTCCC TCCCTGAGAGT GAAGGTGGCA TTGCTCTCGG GGAGCCTGGG CTGCGGGTCT 4380
 TTCCCGGAAG CCCCTGGACC CAAGGCCCCG TTGGCCCCC TGGAAAGAAA GGAGAAAAG 4440
 GTGACTCTGA TCCAAAGGCT CCAGGCTTCC CAGGACAAAC TGGGTCTCCG GGTGAGCAGG 4500
 GCCCAGCGGG ACCTCTCTGA GCTATTGGCC CCAAAGGTGA CCGGGGCTTT CCAGGGCCCC 4560
 TGGGTGAGGC TGAGAGAAAG GCGGAACGTG GACCCCCAGG CCCAGCGGGA TCCCGGGGGC 4620
 TGCCAGGGGT TCGTGGAGCT CCTGGAGCCA AGGGTCTCGA AGGGCCACCA GGACCCACTG 4680
 GCGCGCAAGG AGAGAAAGGG GAGCCTGGTC GGCCTGGGGA CCCTGCACTG GTGGGACCTG 4740
 CTGTGTCTGG ACCCAAAGGA AAAAAGGGAG ATGTGGGGCC CGCTGGGGCC AGAGGAGCTA 4800
 CCGGAGTCCA AGGGGAACGG GGCACCCCG GCTTGGTTCT TCCTGGAGAC CCTGGCCCCA 4860
 AGGGAGACCC TGGAGACCGG GGTCCCATTT GCCTTACTGG CAGAGCAGGA CCCCAGGTG 4920
 ACTCAGGGCT TCCCTGGAGAG AAGGGAGACC CTGGCGGGCC TGGCCCCCA GACCTCTGTT 4980
 GCCCCCGAGG ACCAGATGGT GAAGTTGGAG AGAAAGGTGA CGAGGGTCTT CCGGGTGACC 5040
 CGGGTTCTGG TGAAGAGCTC GCGGAGCGTG GCCTTCCGGG GGCACCTGGA GTTCGGGGCC 5100
 CTGTGGGTGA AAAGGGAGAC CAGGGAGATC CTGGAGAGGA TGGACGAAAT GGCAGCCCTG 5160
 GATCATCTGG ACCCAAGGGT GACCGTGGGG AGCCGGGTCC CCCAGGACCC CCGGGAGCGC 5220
 TGGTAGACAC AGGACCTGGA GCCAGAGAGA AGGGAGAGCC TGGGGACCGG GGAACAAGAG 5280
 GTCTCTGAGG GCCCAAGGGT GATCCTGGCC TCCCTGGAGC CCCTGGGGAA AGGGGCAATT 5340
 AAGGGTTTGG GGAACCCCA GGCACCCAGG GGGACCCAGG TGTCCGAGGC CCAGCAGGAG 5400
 AAAAGGGTGA CCGGGTCTCC CCTGGGCTGG ATGGCCGGAG CCGACTGGAT GGGAAACAG 5460
 GAGCGCTGGG GGCCTCTGGG CCGAATGGTG CTGCAGSCAA AGCTGGGGAC CCAGGGAGAG 5520
 ACGGGCTTCC AGGCTCTGCT GGAGAACAG GCCTCTCTGG CCCTCTGTTG CCCCCTGGAT 5580
 TACCGGGAAA GCGAGCGGAG GATGGGAAC CTGGCCTGAA TGGAAAAAAC GGAGAACCTG 5640
 GGAACCTTGG AGAAGACGGG AGGAAGGGAG AGAAAGGAGA TTCAGCGGCC TCTGGGAGAG 5700
 AAGGTCTGTA TGGCCCAAG GGTGAGCGTG GAGCTCTCTG TATCTTGA CCCCAGGGGC 5760
 CTCAGGCTCT CCAAGGCCCA GTGGGCCCTC CTGGCCAGGG TTTTCTGTTG GTCCAGGAG 5820
 GCAAGGGCCC CAGGGGTGAC CAGGGGAGAG CTGGATCCAA AGGGGAGCAG GGCCTCTCTG 5880
 GAGAGCGTGG CCTGGAGAGA GAGCCTGGAA GTGTGCGGAA TGTGGATCGG TTGCTGGAAA 5940
 CTGCTGGCAT CAAGGCATCT GGCCTCGCGG AGATCGTGA GACCTGGAT GAGAGCTCTG 6000
 GTAGCTTCTT GCCTGTGCCC GAACGGCGTC GAGGCCCCAA GGGGACTCA GGCAGACAG 6060
 GGCCTCCAGG CAAGGAGGCG CCATCTGGCT TTCTGGAGA ACGCGGGCTG AAGGGGAGC 6120
 GTGGAGACCC TGGCCTCAG GGGCCACCTG GTCTGGCCCT TGGGGAGAGG GGCCTCCCTG 6180
 GGCCTTCCGG CCTTCCGGGG GAGCCTGGAA AGCCTGGTAT TCCCGGGCTC CCAGCAGGG 6240
 CTGGGGGTGT GGGAGAGCA GGAAGGCCAG GAGAGAGGGG AGAACGGGGA GAGAAAGGAG 6300
 AACGTGGAGA ACAGGGCAGA GATGGCCCTC CTGGACTCCC TGGAAACCCCT GGGCCCCCG 6360
 GACCCCTTGG CCCCAGAGTG TCTGTGGATG AGCCAGGTCC TGGACTCTCT GGAGAACAG 6420
 GACCCCTTGG ACTCAAGGGT GCTAAGGGGG AGCCGGGCGC CAATGGTGAC CAAGGTCCCA 6480
 AAGGAGACAG GGGTGTGCCA GGCATCAAAG GAGACCGGGG AGAGCCTGGA CCGAGGGGTG 6540
 AGGACGGCAA CCGGGTCTA CCAGGAGAGC GTGGTATGGC TGGGCTGAA GGGAGCCCG 6600
 GTCTGAGGG GTCAAGAGGC CCCCCTGGCC CAGTGGGTGG TCATGGAGAC CCTGGACCAC 6660
 CTGGTCCCC GGGTCTTGT GGCCTGACG GACCCCAAGG ACCTTCTGGC CTGAAGGGGG 6720
 AGCCTGGAGA GACAGGACCT CCAGGACGGG GCCTGACTGG ACCTACTGGA GCTGTGGGAC 6780
 TTCTGGACC CCCCCTCTT TCGGCTTCA GGGGTCTCCA GGTTTGCTCTG 6840
 GACAAGTGGG GSAGACAGGG AAGCCGGGAG CCCCAGGTCC AGATGGTGCC AGTGGAAAAG 6900
 ATGGAGACAG AGGGAGCCCT GGTGTGCCAG GGTCAACAGG TCTGCTTGGC CCTGTGGAC 6960
 CTAAGGGAGA ACCTGGCCCC ACGGGGGCCC CTGGACAGGC TGTGGTGGG CTCCCTGGAG 7020
 CAAAGGGAGA GAGAGGAGCC CTTGGAGGCC TTGCTGGAGA CTTGGTGGGT GAGCCGGGAG 7080
 CCAAGGGTGA CCGAGGACTG CAGGGGCGG GAGGCGAGAA GGGTGAAGCT GGCCTGACG 7140
 GGGAGCCCGG AGACCTTGGG GAAGATGGTC AGAAAGGGGC TCCAGGACCC AAAGGTTTCA 7200
 AGGGTGAACC AGGATCGGG GTCCCGGGCT CCCCCTGGCC TCTGGGCCCT CCAGGTGTGA 7260
 AGGGAGATCT GGGCCTCTCT GGCCTGCCCG GTGCTCTTGG TGTGTGGG TTCCCGGGTC 7320
 AGAGGGCCCC TCGAGGAGAG ATGGGTGAGC CAGGGCCTAG TGGAGAGCGG GGTCTGGCAG 7380
 GCGCCCCAGG GAGAGAGGA ATCCAGGAC CCCTGGGGCC ACCTGGACCA CCGGGGTGAG 7440
 TGGAGCACCT TGGGGCTCT GGACTCAAAG GAGACAAGGG AGACCTGGA GTAGGGCTGC 7500

	CTGGGCCCCG	AGGCGAGCGT	GGGAGGCCAG	GCATCCGGGG	TGAAGATGGC	CGCCCCGGCC	7560
	AGGAGGAGCC	CCGAGGACTC	ACGGGGCCCC	CTGGCAGCAG	GGGAGAGCGT	GGGGAGAAAG	7620
	GTGATGTTGG	GAGTGACAGG	CTAAGGGGTG	ACAAGGGAGA	CTCAGCTGTG	ATCCTGGGGC	7680
5	CTCCAGGCCC	ACGGGGTGCC	AAGGGGGACA	TGGGTGAACG	AGGGCCTCGG	GGCTTGGATG	7740
	GTGACAAAGG	ACCTCGGGGA	GACAATGGGG	ACCTGTGTGA	CAAGGGCAGC	AAGGGAGAGC	7800
	CTGGTGACAA	GGGCTCAGCC	GGGTTGCCAG	GACTGCGTGG	ACTCCTGGGA	CCCCAGGGTC	7860
	AACTCGGTGC	AGCAAGGATC	CCTGGTGACC	CGGGATCCCC	AGGAAAGGAT	GGAGTGCCTG	7920
	GTATCCGAGG	AGAAAAGGA	GATGTTGGCT	TCATGGGTCC	CGGGGCGCTC	AAGGGTGAAC	7980
10	GGGGAGTGAA	GGGAGCCTGT	GGCCTTGATG	GAGAGAAAGG	AGACAAGGGA	GAAGCTGGTC	8040
	CCCCAGGCCG	CCCCGGGCTG	GCAGGACACA	AAGGAGAGAT	GGGGAGCCTT	GGTGTGCCGG	8100
	GCCAGTCCGG	GGCCCCCTGG	AAGGAGGGCT	TGATCGGTCC	CAAGGGTGAC	CGAGGCTTTG	8160
	ACGGGCAGCC	AGGCCCAAG	GGTGACCAGG	GCGAGAAAGG	GGAGCGGGGA	ACCCAGGAA	8220
	TTGGGGCTTT	CCCAGGCCCC	AGTGGAAATG	ATGGCTCTGC	TGGTCCCCCA	GGGCCACCTG	8280
15	GCACTGTTGG	TCCCAGAGCG	CCCGAAGGAC	TTCAAGGCCA	GAAGGTGAG	CGAGGTCCCC	8340
	CCGGAGAGAG	AGTGGTGGGG	GCTCCTGGGG	TCCCTGGAGC	TCCTGGCGAG	AGAGGGGAGC	8400
	AGGGGGCGCC	AGGGCCTGCC	GGTCTCGAG	GCGAGAAAGG	AGAAGCTGCA	CTGACGAGG	8460
	ATGACATCCG	GGGCTTTTGT	CGCCAAGAGA	TGAGTCAGCA	CTGTGCTCTG	CAGGGCCAGT	8520
	TCATCGCATC	TGGATCAACA	CCCCCTCCCTA	GTTATGCTGC	AGACACTGCC	GGCTCCACAG	8580
20	TCCATGCTGT	GCTGTGCTGT	CGCGTCTCTC	ATGCAGAGGA	GGAAGAGCGG	GTACCCCTGT	8640
	AGGATGATGA	GTACTCTGAA	TACTCCGAGT	ATTCTGTGGA	GGAGTACCAG	GACCTGAAG	8700
	CTCCTTGGGA	TAGTGATGAC	CCCTGTTCCC	TGCCACTGGA	TGAGGGCTCC	TGCACTGCCT	8760
	ACACCTCTGG	CTGTTACCAT	CGGGCTGTGA	CAGGCAGCAC	AGAGGCGCTG	CACCTTTTGT	8820
	TCTATGGTGG	CTGTGGAGGG	AATGCCAACC	GTTTGGGGAC	CGGTGAGGCC	TGCGAGCGCC	8880
25	GCTGCCACCC	CCGGGTGGTC	CAGAGCCAGG	GGACAGGTAC	TGCCCAGGAC	TGAGGCCAG	8940
	ATAATGAGCT	GAGATTCAGC	ATCCCTTGA	GGAGTCGGGG	TCTCAGCAGA	ACCCCACTGT	9000
	CCCTCCCTCT	GGTGTAGAG	GCTTGTGTGC	ACGTGAGCGT	GCGAGTGAC	GTCCGTATT	9060
	TCACTGACTT	GCTCCGCTGG	GCTTAGCCTT	CCCCCTGTG	GACAAACCCC	CATTGTGGCT	9120
	CCTGCCACCC	TGGCAGATGA	CTACTGTGG	GGGGTGGCT	GTGGCAGTG	AGCGGATGTG	9180
30	ACTGGCGTCT	GACCCGCCCC	TTGACCAAG	CCTGTGATGA	CATGGTCTG	ATTCTGGGG	9240
	GCATTAAAGC	TGCTGTTTAA	AAAGGCAAAA	AA			

Seq ID NO: 63 Protein sequence:
Protein Accession #: NP_000085.1

35	1	11	21	31	41	51	
	MTLRLLVAAL	CAGILAEAPR	VRAQHERERV	CTRLYAADIV	FLLDGSSSIG	RSNFREVRSP	60
	LEGLVLPPSG	AASQAQVRF	TVQYSDDPRT	EFGLDALGSG	GDVIRAIREL	SVKGNTRTG	120
40	AAILHVADEV	FLPLQARPGV	PKVCILITDG	KSQDLVDTA	QLKGGQVKL	FAVGKINADP	180
	BEKLRVASPT	TSDFPFVND	FSILRTLPL	VSRVCTTAG	GVPVTRPPDD	STSAPRDLVL	240
	SEPSQSLVR	QWTAASGPVT	GKVKQYTLPT	GLGQPLPSER	QEVNVPAGET	SVRLRGLRPL	300
	TEQVTIVIAL	YANSIGEAUS	GCTARTALEG	PELTIONTTA	HSLLVANRSV	PGATGYRVTN	360
	RVLSSGPTQQ	QELGPGQGSV	LLRDLEPGTD	YEVTVSTLFG	RSVGPATSLM	ARTDASVEQT	420
45	LRPVILGPTS	ILLSNLVEPE	ARGYRLEWRR	ETGLEPPQKV	VLPSPDVTRYQ	LDGLQPGTEY	480
	RLTLTYLLEG	HEVATPATVV	PTGPELPVSP	VTDLQATELF	GQRVVRVSWSP	VPGATQYRII	540
	VRSTQGVERT	LVLPGSQATF	DLDDVQAGLS	YTVRVSAVVG	PREGSASVLT	VRRPETPLA	600
	VPLRLRVVSD	ATRRVWAGP	VPASGFRIS	WSTGSGPESS	QTLPPDSTAT	DITGLQPGTT	660
	YQVAVSVLRG	REEGPAIVIV	ARTDPLGPFV	TVHVTQASSS	SVTITWTRVP	GATGYRVSMH	720
50	SAHGPEKSLQ	VSGEATVAEL	DGLEPDTEYT	VHVRHVAAGV	DGPPASVVR	TAPEPVGRVS	780
	RLQILNASSD	TLRITWVGT	GATAYRLAWG	RSEGGPMRHQ	ILPGNTDSAE	IRGLEGGVSY	840
	SVRVTLVGD	REGTPVSIIV	TTPEAPPAL	GTLHVVRQGE	HSLRLRWEFV	PRAQGLLHW	900
	QPGGQGEQSR	LVGPELSSYH	LDGLEPATQY	RVLRSVLGPA	GEGSPAEVTA	RTESPRVPSI	960
	ELRVVDTSID	SVTLANTFVS	RASSYILSWR	PLRGPGQEV	GSPQTLPGIS	SBQRVTGLEP	1020
55	GVSYIFSLPT	VLGGVRGPEA	SVTQTIVCPR	GLADVVFLPH	ATQDNHRAE	ATRRVLERLV	1080
	LALGPLGPPQ	VQGLPLSYSH	RPSPLFPLNG	SHDLGIIILQ	IRDMFYMDDP	GNNLTAVVT	1140
	AHRYMLADPA	PGRRQHVPGV	MVLLVDEPLR	GDIFSPIREA	QASGLNVVML	GMAGADPEQL	1200
	RRLAPGQSV	QTFPAVDGPF	SLDQAVSGLA	TALCQASFTT	QPRPEPCFVY	CPKQKGEKPG	1260
	EMGLRGQVGP	PGDPGLPGRT	GAPGPQGGPG	SATAKGERGF	PGADGRPGSP	GRAGNPGTGP	1320
60	APGLKSGPGL	PGPRGDPGER	GPRGPKGEPG	APGQVIGGEG	PGLPGRKGD	PGSPGPGPRG	1380
	PLGDPGPRGP	PGLPFTAMKG	DKGDRGERGP	PGPGEGGIAF	GEPLGLPLGE	SPGPQGVVGP	1440
	PGKKGKGGD	EDGAPGLPGQ	PGSPGEGQPR	GPPGAIGPKG	DRGFPGLPGE	AGEKGERGPP	1500
	GPAGSRGLPG	VACRPGAKGP	EGPPGPTGRQ	GEKGEPRPG	DPAVVGPAVA	GPKGEGKDVG	1560
	PAGPRGATGV	QGERGPPGLV	LPGDPGPKGD	PGDRGPIGLT	GRAGPPGDSG	PFGEKGDPRG	1620
65	PGPPGVPVGR	GRDGEVGEKG	DEGPPGDPGL	PGKAGERGLR	GAPGVRGPGV	EKGQDQDPGE	1680
	DGRNGSPGSS	GPKGDRGEPG	PPGPPGRLVD	TGPGAREKGE	PGDRGQEGPR	GPKGDPGLPG	1740
	APGERGIEGF	RGPFGQGGDP	GVRGPAGEKG	DRGPPGLDGR	SGLDGKPGAA	GSPGNGAAG	1800
	KAGDPGRDGL	PLRGRBQGLP	GPSGPPGLPG	KPGEDGKPLG	NKNGEPGDP	GEDGRKGEKG	1860
70	DSGASGREGR	DGPKGERGAP	GILGPQGGPG	LPGVPVPGPG	GFPVVPGGTG	PKGDRGETGS	1920
	KGEQGLPGER	GLRGEPSGVP	NVDRLLTAG	IKASALREIV	ETWDESSGTF	LPVPERRRGP	1980
	KGDSGEQGGP	GKEGPIGFPG	ERGLKGRDGD	PGPQCGPGLA	LGERGPPGPS	GLAGEPKKPG	2040
	IPGLPGRAGG	VGEAGRPERG	GERGEKGERG	EQGRDGPPLG	PPTPGPPGPP	GPKVSVDEPG	2100
	PGLSGEQGPP	GLKQAKGEPG	SNGDQGPFGD	RGVPGIKGDR	GEPGPRQDQ	NPGLPGERGM	2160
	AGPEGKPLGQ	GPRGPPGPGV	GHGDPGPPGA	PGLAGPAGPQ	GPSGLKGEPE	ETGPPGRGLT	2220
75	GPTGAVGLPG	PPGFSGLVGP	QGSPLPGQV	GETGKPGAPG	RDGASGKGDG	RGSPPGVGSP	2280
	GLPGVPVPGK	EPGFTGAPGQ	AVVGLPGAKG	EKGAPGGLAG	DLVGEPGAKG	DRGLPGPRGE	2340
	KGEAGRAGEP	CDPGEQGXG	APGPKGEKGD	PGVGVPGSPG	PPGPPGVKGD	LGLPGLPGAP	2400
	GVVGFPGQTG	FRGEMQPGP	SGERGLAGPP	GREGIPGPLG	PPGPPGVSVP	PGASGLKGDG	2460
	GDPGVGLPGP	RGERGEPGIR	GEDGRPGQEG	PRGLTGPPGS	RGERGERGDP	GSAGLKGDGK	2520
80	DSAVILPFGP	PRGAKDMGE	RGRGLDGDG	GPRGDNPDG	DKSGKEPFGD	KGSAGLPGLR	2580
	GLLGPQGGPG	AAGIGLPGGS	PKDGVPGIR	GEKGDVFGIR	PRGLKGERGV	KGACGLDGEK	2640
	GDKGEAGPPG	RPLAGHKGE	MGEPPVPGQS	GAPKGEGLIG	PKGDRGFDGQ	PGPKGDQGEK	2700
	GERGTPGICG	PPGFSNGDGS	AGPPGPPGVS	GPRGPEGLQG	QKGERGPPGE	RVVGAPGVPG	2760
	APGERGEQGR	PGPAGPRGKG	GEAALTEDDI	RGFVRQEMSQ	HCACQGGFIA	SGSRPLPSYA	2820
85	ADTAGSLQHA	VPVLRVSHAE	EZERVPFEDD	EYSEYSEYSV	EYQDPEAPW	DSDDPCSPL	2880
	DEGSTCTAYT	RWYHRAVTGS	TEACHPFVYG	GCGGNANRFG	TREACERRCF	PRVVSQSGTG	2940
	TAQD						

5
1 11 21 31 41 51
ATGTCCTATC AACAGCAGCA GTGCAAGCAG CCTGCCAGC CACCTCTGT GTGCCCCAG 60
CCAAAGTGCC CAGAGCCATG TCCACCCCG AAGTGCCCTG AGCCCTGCCC ACCACCAAG 120
10 TGTCCACAG CCGTCCCAAC TCAGCAGTGC CAGCAGAAAT ATCTCTCTGT GACACCTTCC 180
CGACCTGCC AGCCAAAGTA TCCACCGAAG AGCAAGTAA

15
1 11 21 31 41 51
MSYQQQQCKQ PCQPPFVCPT PKCPFCPPF KCPEFCPPK CPQFCPPQC QQKYPFVTPS 60
PPCQPKYPPK SK

	1	11	21	31	41	51	
25	TAGTCGGAGC	GAGGTGGCGA	GTGCTGTGAGC	CCGCGCGCGC	CCCGAGAGCG	GCTGTCAGCCG	60
	CGCGCCGGCG	GAAAGGAGAGG	GCGAGGGCGCG	CCCGAGCGCG	CGCGCCGGCG	GCGACCGCGCG	120
	CGCGCCGGCG	CGACGCCAAC	GAGGTGCGCG	CGAGCGCGCG	CAGCTCTCCG	CGCCGCCCGCG	180
	CGGGCGGGGG	GGCGCGGGCC	ACAGGCCCCCT	GCTCCGCGCG	TGGTTTGCAG	ACCGCGGGCG	240
30	CGGATGTGCG	CGCGCGCGCG	TTAGAGTAGT	GCTTGCGGTG	GGCGAGGAGC	CGCGCGAGCC	300
	GCGCGCGCGC	GAGCGCGCGG	CAGGAGCTCT	GGGAGCGCGC	CGCGCGCGCG	CGCGCGCGCG	360
	CGCGGGCCCC	GACGCGCCCC	CGCGCGCCCC	CGCCGCCCGA	CACACATGAG	ATTCTTCAGG	420
	CTCATTTTCA	AGTGCTTCGT	GGACTGCTCT	TGACTGCGCC	CGCGCGCGCG	CGACCGCGCG	480
	CGTCCGCGCG	CGGCCCCGTC	CCCCGGCGCG	GGCGCCCCCG	GGCCCCCGCG	CGGCCCCGCG	540
35	CTCGGGGGCG	CTCCCCGGGT	CGCGCGGTGC	CCCCCGCGCT	ACCGCGCGCC	CGCGTGGAGC	600
	GCGCGACACC	CGCGCGCGCG	GTGCGCGGCT	CGCGGGCCAT	GGCGAAGAG	AGCGCGGAGA	660
	AGGCGACTTA	TAGGCTGTCT	GGCGACAGAA	AGAGAGGGCC	CCCTCATCGG	CGCGGGCGCG	720
	ACGGGGCCCC	GGCGAAGGGC	GACGGCCCCG	TGGGCTTGGG	GACACCGCGG	GGCGCGCTCG	780
40	CGGTGCGCGC	CGCGCGCGCG	TGGACGCGCC	AGATGGACTT	CATCATGTGG	TGGCTGGGCT	840
	TGCGCGTGGG	CTGGGGCAAC	GTGTGGGAGT	TCCCTTACCT	GTGCTACAG	AACGGCGGAG	900
	GTGTGTTTCT	TATTTCCCTAC	GTGCTGATCG	CGCTGGTTGG	AGGAATCCCC	ATTTTCTTCT	960
	TAGAGATCTC	CTGGGGCGAG	TTCATGAAAG	CGCGCAGCAT	CATGTCTCTG	AACACTCTGTC	1020
	CCCTGTTTCA	AGGCTCTGGG	TACGCTTCCA	TGGTGTCTGT	CTTCTACTCG	AACACTTACT	1080
45	ACATCATGGT	GCTGGCTGCC	GGCTTCTATT	ACCTGGTCAA	GTCTTTTACC	ACACGCTCTC	1140
	CTCGGGCCAC	ATGTGGCCAC	ACCTGGAAAC	CTCCGACTCT	GTGGAGATCT	GTCTCCGATC	1200
	AGAGACTGTG	CATGTCCAGC	CTGGCCAAAC	TCACTCTGTG	CCAGCTTGCT	GACCGCGCGT	1260
	CCCTCTGTCT	CGAGTCTCTG	GAGAAACAAAG	GTCTGAGGCT	GTCTGGGGGA	CTGGAGTCTC	1320
	CAGGGCGCCT	CACCTGGGAG	GTGACCTCTT	GCTCTGTGCG	GTCTGGGTGT	CTGTGTTACT	1380
	TCGTGTCTGT	GAAGGGGGTC	AAATCCACCG	GAAAGATCTG	GTACTTCACT	GCTCAATTCC	1440
50	CTCATGTGGT	CTCGGTGGTG	GTGCTGGTGC	GTGGAGTGTG	GTGCTCTGCG	GGCCGTGATG	1500
	GCATCATTTA	CTATCTCAAG	CCTGACTGGT	CAAAGCTGGG	GTCCCTCTAG	GTGTGGATAG	1560
	ATGCGGGGAG	CCGATTTTTC	TTTTTCTACG	CCATTGGCTC	GGGGCGCCCT	ACAGCGCCCTG	1620
	GCAGCTACAA	CGAGCTTAAAC	AACAACTGCT	CAAGAGACGC	CATCATCTCT	CTCTCTATCA	1680
	ACAGTGGGAC	CAGCTTCTTT	CTGGCTTCTG	TGGTCTTCTC	CATCTGGGCG	TTCTATGGCT	1740
55	CAGAGCAGGG	GTGCATCAAT	TCCAAGGTGG	CAGAGTCACT	CGCGGGCGCT	GGCTTCTATC	1800
	CCTACCGCGG	GGCTGTCAAG	CTGATGCCAG	TGGCCCCACT	CTGGGCTGCC	CTGTCTTCTT	1860
	TCATGCTTGT	GCTCGTCTGGT	CTGACAGGCC	AGTTTGTAGG	TGTGGAGGGC	TTCTATACCG	1920
	GGCTCTCTGA	CTCTCCCGCG	GGCTCTTACT	ACTCTCGGTT	CCAAAGGGAG	ATCTCTGGTG	1980
	CCCTCTGTGT	TGCCCTCTGC	TTTGTCTATC	ATCTCTCCAT	GGGACTGATG	GGCGGGATGT	2040
60	ACGTTCTTCA	CTGTCTTGAG	TACTACTCGG	CACAGCGGCG	CACCTCGTCT	TGGCAGAGCT	2100
	TTTGGAGTGT	CGTGGTGGTG	GGCTGGGTGT	ACGGAGCTGA	CGGCTTCACT	GACGACATTT	2160
	CTGTATGATG	CGGGTACCGA	CTTCTGCCCT	GGATGAAATG	GTGCTGTGTC	TTCTTTCACC	2220
	CGCTGGTCTG	CATGGGCACT	TTCAATTTCA	GGTCTTGTTA	CTACGAGCGG	CTGTGTTACA	2280
65	ACAACACCTA	CGTGTACCGG	TGGTGGGGTG	AGGGCACTGG	CTGGGCGCTC	GGCGTGTCTC	2340
	CATGCTGTGT	CTGCGCGCTG	CACCTCTGCG	GCTCGCTCTC	CAGGGCCAGG	GGCCACTATG	2400
	CTGAGCGCTG	CGACGACCTT	ACCCAGGCCA	TCGTGGGCGT	CCACCACTTG	GAGTACCGAG	2460
	CTCAGGACGC	AGATGTCTAG	GGCCTGACCA	CCCTGACCCC	AGTGTCCGAG	AGCAGCAAGT	2520
	TGCTGTGGTG	GGAGAGTGTG	ATGTGACAAC	TACGCTCACA	TACCAAGCTC	ACCTCTGGTA	2580
70	GCCATAGCAG	CCCTCTGCTT	AGGCCGCCCT	CACCCCTCCA	GGGGCGCTGC	CTTTCCCTGA	2640
	CATTTTGGGG	CTGCTCTCTG	GGAGGAGAGG	GAGAAAGGAC	CATGAGTGTG	CTATAAAACA	2700
	ACTTTTTCCTA	TTTTTATAAA	AAGGCCAAAA	ATATACCAAC	CCACCAAAAA	TAGATGGCTC	2760
	TCCCTCTCCA	GGCTTAGCGG	AGCTGTGCTT	AGGGCCCGCC	TAGTGCCCCA	CCCCCAACCA	2820
	CAGTGTGCTA	CTCTCTCTGC	CCCTGCCACG	CCCACCCCCT	GGCCACCTCT	CCAGGCTCTG	2880
	CTCTGACGTA	CAACCGGTGG	TGACCCCTCA	CCGCCAGGAG	AGCAGTGGCA	TCTTGGGAAA	2940
75	TGTGAGGAAG	GGAAGAGAGG	AGAGACGGGA	GGGAGGAGAG	GAGAGGAGAG	GAGGCGACGG	3000
	GAGGGGCGAG	AGAACAAGG					

CTTCTGTGTA GCAGCTTTAA CCCACGTTTG TCTGTCACGT CCAGTCCCGA GACGGCTGAG 3780
 TGACCCCAAG AAGGGCTTCC CCGACACCCA GACAGAGGCT GCAGGGCTGG GGCTGGGTGA 3840
 GGGTGGCGGG CCTGCGGGGA CATCTACTG TGCTAAAAAG CCACTGCAGA CATAGCAATA 3900
 AAAACATGTC ATTTTCC

Seq ID NO: 67 Protein sequence:
 Protein Accession #: NP_005620.1

1 11 21 31 41 51
 10 | | | | |
 MAKSAENGI YSVSGDEKRG PLIAPGPDGA PAKGDGPVGL GTPGGR LAVP PRETWTRQMD 60
 PIMSCVGFPAV GLGNVVRFFY LCYKNGGGVF LIPYVLIALV GGIPIFFLEI SLGQPMKAGS 120
 INVMNICPLF KGLGYASPMVI VFYCNYYIIM VLAWGFFYLK KSFTTTLFWA TCGHTWNTPD 180
 15 CVEIFRHEDC ANASLANLTC DQLADRRSPV IEFWENKVLRL LSGGLEVPGA LNWEVTLCLL 240
 ACWVLVYFCV WKGVKSTGKI VYPTATFPYV VLVVLLVRGV LLPGALDGIY YYLKPDMWSKL 300
 GSPQVWIDAG TQIFPFAIYG LGALTALGSY NRPNNNCKYD AIILALINSQ TSFFPAGPVVF 360
 SILGFMAAEQ GVHISKVAES GPGLAFLIAYP RAVTLMFPVAP LWAALFFPML LLLGLDSQFV 420
 GVEGFTIGLL DLLPASYVYFR FQREISVALC CALCFVIDLS MVTGGMYVP QLPDYYSASG 480
 20 TLLHQAQFWE CVVVAWVYGA DRFMDDIACM IGYRCPWPMK WCNWFFTPLV CMGIFIPNVV 540
 YVEPLVYNNY YVYPWNGEAM GWAFALSSML CVPLHLGLCL LRAGTMAER WQHLTQPIWG 600
 LHHLEYRAQD ADVRGLTTLT PVSESSKVVV VESVM

Seq ID NO: 68 DNA sequence
 Nucleic Acid Accession #: NM_021953.1
 Coding sequence: 178-2469

1 11 21 31 41 51
 30 | | | | |
 GGCACGAGGG GGACCCGGCC GGTCCGGCGC GAGCCCCCGT COGGGGCCCT GGCTCGGCC 60
 CAGAGGTGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG GCGGCGACTG 120
 CAGTCTGGAG GGTCCCACTT TGTGATTCTC AATGGAGAGT GAAACGCGAG ATTCATAATG 180
 AAAGCTAGCC CCGCTCGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCTT TCCTGTTCAA 240
 35 AATGCCCCAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCGCTGCCCA ACAGAGAGTCT 300
 AATCAAGCAG AGGCCTCCAA GGAAGTGGCG GAGTCCAACT CTTGCAAGTT TCCAGCTGGG 360
 ATCAAGATTA TTAACACACC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCACCAAT 420
 GCTAATATTG ACAGCATCAT CACAGCACTG ACTGCCAAGG GAAAGAGAGT TGGCAGTAGT 480
 GGGCCCAACA AATTCATCCT CATCAGCTGT GGGGGAGCCC CAACTCAGCC TCCAGGACTC 540
 40 CGGCCTCAAA CCCAAACAG CTATGATGCC AAAAGGACAG AAGTGACCCT GGAGACCTTG 600
 GGACCAAAAC GATGCTAGT GGAATGTGAAT CTTCTAGAC CACCTGGAGC CCTTTGCGAG 660
 CAGAAACGGG AGACCTGTGC AGATGGTGAG GCACGAGGCT GCACATATCA CAATAGCCTA 720
 TCCAACTATC AGTGGCTTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780
 CAAGAGATGG AGGAAAAGGA GAATTGTGAC CTGGAGCAGC GACAGGTTAA GGTGAGGAG 840
 45 CTTTCGAGAC CATCAGCTGC CTGGCAGAAC TCTGTGTCTG AGCGGCCACC CTACTCTTAC 900
 ATGGCCATGA TACAATTGCG CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
 ATCTATACGT GGAATGAGGA CCACTTTCCC TACTTTAAGC ACATTTGCCA GCCAGGCTGG 1020
 AAGAACTCCA TCCGCCACAA CCTTTCCCTG CACGACATGT TTGTCCGGGA GACGCTGCGC 1080
 AATGGCAAGG TCTCCTTCTG GACCATTGAC CCCAGTGCCA ACCGCTACTT GACATTGGAC 1140
 50 CAGGTGTTTA AGCCACTGGA CCCAGGGTCT CCACAATTGC CGAGGACATT GGAATCACAG 1200
 CAGAAACGAC CGAATCCAGA GCTCCGCGCG AACATGACCA TCAAAACCGA ACTCCCCCTG 1260
 GGGCAGCGCG GGAAGATGAA GCCACTGCTA CCACGGGTCA GCTCATACCT GGTACCTATC 1320
 CAGTTCOCGG TGAACCACTG ACTGGTGTG CAGCCCTCGG TGAAGGTGCC ATTTGCCCTG 1380
 GCGGCTTCCC TCATGAGCTC AGAGCTTGCC CGCCATAGCA AGCGAGTCCG CATTGCCCCC 1440
 AAGGTGTGCT TAGCTGAGGA GGGGATAGCT CCTCTTTCTT CTGCAGGACC AGGGAAAGAG 1500
 55 GAGAAACTCC TGTTTGGAGA AGGGTTTCTC CTTTGTCTTC CAGTTTCAGC TATCAAGGAG 1560
 GAAGAAATCC AGCCTGGGGA GGAATGCCA CACTTAGCGA GACCCATCAA AGTGGAGAGC 1620
 CCTCCCTTGG AAGAGTGGCC CTCGCCGCCC CCATCTTTCA AAGAGGAATC ATCTCACTCC 1680
 TGGGAGGATT CGTCCCAATC TCCCACCCCA AGACCCCAAG AGTCTACAG TGGGCTTAGG 1740
 60 TCCCAACCCC GGTGTGCTC GGAATGCTT GTGATTCAAC ACAGGGAGAG GAGGGAGAGG 1800
 AGCCGCTCTC GGAAGAAACA GCATCTACTG CCTCCCTGTG TGGATGAGCC GGAGCTGCTC 1860
 TTCTCAGAGG GGCCAGTAC TTCCCGCTGG GCCCGAGAGC TCCGTTCCCG AGCAGACTCC 1920
 TCTGACCTTG CCTCCAGCT CAGCTACTCC CAGGAAGTGG GAGGACCTTT TAAGACACCC 1980
 ATTAAGGAAA CGCTGCCATC CTCTCCACC CGAGCAAAAT CTGTCTCCC CAGAACCCCT 2040
 65 GAATCTCGA GGTCTACGCC CCCAGCCAAA GTAGGGGGAC TGGATTTCAG CCGAGTACAA 2100
 ACCTCCAGG GTGCTCTGA CCCCTTGCCCT GACCCCTGGG GGCTGATGGA TCTCAGCACC 2160
 ACTCCCTTGC AAGATGCTCC CCCCTTGAA TCACCGCAA GGTCTCTCAG TTCAGAACCC 2220
 TTAGACCTCA TCTCCGTCCT CTTTGGCAAC TCTTCTCCCT CAGATATAGA CGTCCCCAAG 2280
 CCAGGCTCCC CGGAGCCACA GGTTCCTGGC CTTGCAGCCA ATCGTCTCT GACAGAAGGC 2340
 70 CTGGTCTGG ACACAATGAA TGACAGCCTC AGCAAGATCC TGCTGGACAT CAGCTTCTCT 2400
 GGCTGGAGG GAGACCCACT GGGCCCTGAC AACATCAACT GGTCCAGTT TATTCTTGAG 2460
 CTACAGTAGA GCCTGCCCT TGCCCTGTG CTCAAGCTGT CCACCATCCC GGGCACTCCA 2520
 AGGCTCAGT CACCCCAAGC CTCGAGTGA GGACAGCAGG CAGGGACTGT TCTGCTCCTC 2580
 ATAGCTCCCT GCTGCCTGAT TATGCAAAAG TAGCAGTCAC ACCCTAGCCA CTGCTGGGAC 2640
 75 CTTGTGTTCC CCAAGAGTAT CTGATCTCTC TGCTGTCCCT GCCAGGAGCT GAAGGGTGGG 2700
 AACACAAGAG GCAATGGTGA AAAGAGATTA GGAACCCCCC AGCCTGTGTC CATTCTCTGC 2760
 CCAGCAGTCT CTTACCTTCC CTGATCTTTG CAGGGTGGTC CGTGTAATA GTATAAATTC 2820
 TCCAAATTAT CCTCTAATTA TAAATGTAAG CTTATTTCTT TAGATCATTA TCCAGAGACT 2880
 GCCAGAAGGT GGTGAGGATG ACCTGGGGTT TCAATTGACT TCTGTTCTCT GCTTTTAGTT 2940
 80 TTGATAGAG GGAAGACCTG CAGTGACCGG TTTCTTCCAG GCTGAGGTAC CTGATCTTGT 3000
 GGTCTCTCAC TGCAGGGACC CAGACAAGTG GATCTGCTTG CCAGAGTCTT TTTTGCCCTT 3060
 CCTGCGCACC TCCCGGTGTT TCCAAGTCAG CTTTCTCTGCA AGAAGAAATC CTGGTTAAAA 3120
 AAGTCTTTTG TATTTGGTCA GGAGTTGAAT TTGGGGTGGG AGGATGGATG CAACTGAAGC 3180
 AGAGTGTGGG TGGCCAGATG TGGCTATTA GATGTTCTC TGATAATGTC CCAATCATA 3240
 85 CCAGGGAGAC TGGCATTTGAC GAGAACTCAG GTGGAGGCTT GAGAGGCGG AAGGGGCCCC 3300
 TGACCTGCCT GGTCTCTTGA GCTTGGCCCT CAGCTTTGCA AAGAGCCACC CTAGGCCCCA 3360
 GCTGACCGCA TGGGTGTGAG CCAGCTTGAG AACACTAACT ACTCAATAAA AGCGAAGGTG 3420
 GACCNAAAAA AAAAAAAAAA AAAA

Seq ID NO: 69 Protein sequence:
Protein Accession #: NP_068772.1

```

5      1      11      21      31      41      51
|      |      |      |      |      |
MKASPRRLPI LKRRRLPLPV QNAPSETSEE EPKRSPAQQE SNQAEASKEV AESNSCKFFPA 60
GIKILNHPTM PNTQVVAIFN NANIHSIITA LTAKGRESGS SGPNKFILIS CGGAPTQPPG 120
LRPQTQTSYD AKRTEVTLET LGPKPAARDV NLPRPPGALC EQKRETCADG EAAGCTINNS 180
10    LSNIQWLRLM SSDGLGSRSI KQEMEKEKNC HLEQRQVKVE EPSRPSASWQ NSVSERPPYS 240
YMAHIQFAIN STERKRMTLK DIYTWIEDHF PYFKHIAKPG WKNSIRHNL LHMFPVRETS 300
ANGKVSFWPI HPSANRYLTL DQVFKPLDPG SPQLPEHLES QQKRFNPELR RNMTIKTELP 360
LGARRKMKPL LPRVSSYLVP IQFFVNQSLV LQPSVKVPLP LAASLMSSSEL ARHSKRVRIA 420
PKVLLAEBGI APLSSAGPGK EEKLLFGEGF SPLLPVQTIK EEEIQGEEM PHLARPIKVE 480
15    SPLEEWPSF APSPKEESSH SWEDSSQSPT PRPKSYSGL RSPTRCVSEM LVIQHRERRE 540
RSRSRRKQHL LPPCVDEPEL LPSEGPSTSR WAAELPPPAD SSDPASQLSY SQEVGGPFKT 600
PIKETLPISS TPKSKVLPRT PESWRLTPPA KVGGLDFSPV QTSQASDPL PDPLGLMDLS 660
TFLQSAAPPL ESPQRLLSSE PLDLISVFPQ NSSPSDIDVP KPGSPEPQVS GLAANRSLTE 720
20    GLVLDTMNDS LSKILLDISF PGLDEDPLGF DNINWSQFIP ELQ

```

Seq ID NO: 70 DNA sequence
Nucleic Acid Accession #: BC006529.1
Coding sequence: 178-2424

```

25      1      11      21      31      41      51
|      |      |      |      |      |
GGCACGAGGG GGACCCGGCC GGTCCGGGCG GAGCCCCCGT CCGGGGCCCT GGCTCGGCC 60
30    CCAGGTTGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG CGCGCGACTG 120
CAGTCTGGAG GGTCCACACT TGTGATTCTC AATGGAGAGT GAAACGCGAG ATTCATAATG 180
AAAACTAGCC CCGCTGGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCTT TCCTGTTCAA 240
AATGCCCAA GTGAACATCT AGAGGAGGAA CCTAAGAGAT CCCCTGCCCA ACAGGAGTCT 300
AATCAAGCAG AGGCCTCCAA GGAAGTGGCA GAGTCCAAC CTTCGAAGTT TCCAGCTGGG 360
ATCAAGATTA TTAACACACC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCACAAAT 420
35    GCTAATATT ACAGCATCTC CACAGCACTG ACTGCCAAGG GAAAGAGAG TGGCAGTAGT 480
GGGCCCAACA AATTCATCTC CATCAGCTGT GGGGGAGCCC CAACTCAGCC TCCAGGACTC 540
CGGCTCCAAA CCCAAACCCAG CATGATGCC AAAAGGACAG AAGTGACCTT GGAGACCTTG 600
GGACCAAAAC CTGACAGTAG GGATGTGAAT CTTCCTAGAC CACCTGGAGC CCTTTGCGAG 660
CAGAAACGGG AGACCTGTGC AGATGGTGAG GCACGAGGCT GCACTATCAA CAATAGCCTA 720
40    TCCAACATCC AGTGGCTTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780
CAAGAGATGG AGGAAAGGGA GAATTGTCTC CTGGAGCAGC GACAGGTTAA GGTTAGGAGG 840
CCTTCGAGAC CATCAGCGTC CTGGCAGAAC TCTGTGTCGT AGCGGCCACC CTACTCTTAC 900
ATGGCCATGA TACAATTGCG CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
45    ATCTATACGT GGAATGAGGA CCACCTTCCC TACTTTAAGC ACATTGCCAA GCCAGGCTGG 1020
AAGAATCTCA TCGCCACAAA CCTTCCCTCG CAOGACATGT TTGTCCGGGA GACGCTCGCC 1080
AATGGCAGTG TCTCTCTCTG GACCATTCAC CCCAGTGCCA ACCGCTACTT GACATTGGAC 1140
CAGGTGTTTA AGCAGACGAA AGACCGAAT CCAGAGCTCC GCCGGAACAT GACATCAAA 1200
ACCGAATCTC CCTTGGGCGC ACGGCGGAAG ATGAAGCCAC TGCTACCAGC GGTCAAGTCA 1260
50    TACCTGGTAT CTATCCAGTT CCGGCTGAAC CAGTCACTGG TGTTCAGACC CTCGGTGAAG 1320
GTGCCATTGC CCTTGGGCGC TTCCCTCATG AGCTCAGAGC TTGCCCCCA TAGCAAGCGA 1380
GTCCGATGTC CCCCAGAGT GCTGCTAGCT GAGGAGGGGA TAGCTCTCTT TTCTTCTGCA 1440
GGACCAAGGA AAGAGAGAA ACTCTGTTT CGAGAAGGGT TTTCTCTTTT GCTTCCAGTT 1500
CAGACTATCA AGGAGGAAGA AATCCAGCCT GGGGAGGAAA TGCCACACTT AGCGAGACCC 1560
55    ATCAAGATGG AGAGCCCTCC CTGGAAGAG TGGCCCTCCC CGGCCCATC TTTCAAAGAG 1620
GAATCATCTC ACTCTGGGA GGAATCGTCC CAATCTCCCA CCCCAAGACC CAAGAAGTCC 1680
TACAGTGGGC TTAGGTCCCC AACCCGCTGT GTCTCGGAAA TGCTTGATAT TCAACACAGG 1740
GAGAGGAGGG AGAGGAGCCG GTCTCGGAGG AAACAGCATC TACTGCCTCC CTGTGTGGAT 1800
GAGCCGAGAC TGCTCTTCTC AGAGGGGCCC AGTACTTCCC GCTGGGCCCC AGAGCTCCCC 1860
60    TTCCAGCAG ACTCTCTGAA CCCTGCCTCC CAGCTCAGCT ACTCCCAGGA AGTGGGAGGA 1920
CCTTTTAAGA CACCCATTAA GGAACGCTG CCCATCTCCT CCACCCGAG CAAATCTGTC 1980
CTCCCCAGAA CCCCAGATC CTGGAGGCTC ACGCCCCAG CCAAGTAGG GGGACTGGAT 2040
TTCAGCCAG TACAAACCCC CCAGGCTGCC TCTGACCCCT TGCCTGACCC CCTGGGCTG 2100
ATGGATCTCA GCACCACTCC CTGCAAAAGT GCTCCCCCCC TTGAATCACC GCAAGGCTC 2160
65    CTCAGTTCAG AACCTTAGA CCTCATCTCC GTCCCTTTG GCAACTCTC TCCCTCAGAT 2220
ATAGACGTCC CCAAGCCAGG CTCCCGGAG CCACAGGTTT CTGGCCTTGC AGCCAATCGT 2280
TCTCTGACAG AAGGCTCTGT CCTGGACACA ATGAATGACA GCCTCAGCAA GATCCTGCTG 2340
GACATCAGCT TTCTTGGCCT GGAAGAGGAC CCACTGGGCC CTGACAACAT CAATCTGGTC 2400
CAGTTTATTG CTGAGCTACA GTAGAGCCCT GCCCTTGCCC CTGTGCTCAA GCTGTCCACC 2460
70    ATCCCGGGCA CTCGAAGGCT CAGTGCAACC CAAGCCTCTG AGTGAGGACA GCAGGCGAGG 2520
ACTGTTCTGC TCTCATAGC TCCCTGCTGC CTGATTATGC AAAAGTAGCA GTCACACCT 2580
AGCCTATGCT GGGACCTGTG GTTCCCAAG AGTATCTGAT TCTCTGCTG TCCCTGCCAG 2640
GAGCTGAAGG GTGGGAACAA CAAAGGCAAT GGTGAAAGA GATTAGGAAC CCCCGAGCCT 2700
75    GTTTCATTTC TCTGCCAGC AGTCTCTTAC CTTCCCTGAT CTTTGCAGGG TGGTCCGTGT 2760
AAATAGTATA AATTCTCCAA ATTATCTCTT AATTATAAAT GTAAGCTTAT TTCTTAGAT 2820
CATTATCCAG AGACTGCCAG AAGGTGGGTA GGATGACCTG GGGTTTCAAT TGACTTCTGT 2880
TCTCTGCTTT TAGTTTGTAT AGAAGGGAAG ACCTGCAGTG CACGGTTTCT TCCAGGCTGA 2940
GGTACCTGGA TCTTGGGTTT TCACTGCGAG GGACCCAGAC AAGTGGATCT GCTTGCCAGA 3000
80    GTCCCTTTTG CCCCCTCCCT CCACCTCCCC GTGTTTCCAA GTCAGCTTTC CTGCAAGAG 3060
AAATCTCTGT TAAAAAAGTC TTTTGTATTG GGTGAGGAGT TGAATTTGGG GTGGGAGGAT 3120
GGATGCAACT GAAGCAGAGT GTGGGTGCCC AGATGTGCGC TATTAGATGT TTCTCTGATA 3180
ATGTCCCCAA TCATACCAGG GAGACTGGCA TTGACGAGAA CTCAGGTGGA GGCCTGAGAA 3240
GGCCGAAGGG GCCCTGACCC TGCCTGGCTT CCTTAGCTTG CCCCTCAGCT TTGCAAGAG 3300
85    CCACCTTAGG CCCCAGCTGA CCGCATGGGT GTGAGCCAGC TTGAGAACAC TAACTACTCA 3360
ATAAAAAGCA AGGTGAAAAA AAAAAAATA AAAAAA

```

Seq ID NO: 71 Protein sequence:
Protein Accession #: AA06529.1

```

1      11      21      31      41      51
|      |      |      |      |      |
5 MKTSRRRLPI LKRRRLPLFV QNAPSETSEE EPKRSPAQQE SNQAEASKEV AESNSCKFPA 60
GIKIINHPTM PNTQVVAIPN NANIHSIITA LTAKGKESGS GPNKIFILIS CGGAPTQPPG 120
LRPQTQTSYD AKRTEVTLLET LGPKPAARDV NLPRPPGALC EQKRETCADG EAAGCTINNS 180
LSNIQWLKRM SSDGLGSRSI KQEMEKEKNC HLEQRQVKVB EPSRPSASWQ NSVSRPPYS 240
YWAMIQFALN STERRKMTLK DIYTWIEDHF PYFKHIAKPG WNSIIRHNS LHDMFVRETS 300
ANGKVSFWTI HPSANRYLTL DQVFKQKRP NPELRNMTI KTELPLGARR KMKPLLRVS 360
10 SYLVPIQFPV NQSLVLQPSV KVPLPLAASL MSSELRHSHK RVRIAPKVLL AEBGIAPLSS 420
AGPGKEEKLK FEGGFSPLLP VQTIKEEEIQ PGEEMPHLAR PIKVESPPLE EWSPSPAPSK 480
EESHSWEDS SQSPTPRPKK SYSGLRSPTR CVSEMLVIQH RERRERSR SRKQHLPLPCV 540
DEPELLFSEG PSTSRWAEEL PFPADSSDPA SOLSYGQEVG GPFKTPIKET LPISSPTPSK 600
VLPTFPESWR LTTPAKVGLG DPFVQTPQG ASDPLPDPLG LMDLSTPLQ SAPPLESPOR 660
15 LLSSEPLDLI SVFPGNSSPS DIDVPRKPGSP EPQVSGLAAN RSLTEGLVLD TMNDSLKIL 720
LDSFPGLDE DPLGPDNINW SQPIPELQ

```

Seq ID NO: 72 DNA sequence
Nucleic Acid Accession #: U74612.1
Coding sequence: 178-2583

```

1      11      21      31      41      51
|      |      |      |      |      |
25 GGCACGAGGG GGACCCGGCC GGTCCGGCGC GAGCCCCCGT CCGGGGCCCT GGCTCGGCC 60
CCAGGTGGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG GCGGCGACTG 120
CAGTCTGGAG GGTCCACACT TGTGATTCTC AATGGAGAGT GAAAACGCGAG ATTCTAATG 180
AAACTAGCC CCGCTCGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCCT TCCTGTTCAT 240
AATGCCCAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCCCTGCCCA ACAGGAGTCT 300
AATCAAGCAG AGGCCTCCAA GGAAGTGGA GAGTCCAACT CTGCAAGTT TCCAGTGGG 360
ATCAAGATTA TTAACACACC CACCATGCC AACAACGCAAG TAGTGCCAT CCCCACAAAT 420
GCTAATATTA ACAGACTCAT CACAGCACTG ACTGCCAAGG GAAAAGAGAG TGGCAGTAGT 480
GGGCCCAACA AATTCATCTT CATCAGCTGT GGGGGAGCCC CACTCAGCC TCCAGGACTC 540
CGSCCTCAAA CCCAACCCAG CTATGATGCC AAAAGGACAG AAGTGACCTT GGAGACCTTG 600
GGACCAAAAC CTGCAGTAG GATGTGAAT CTTCCTAGAC CACTGGAGC CCTTTCGAG 660
CAGAAACGGG AGACTGTGAG AGATGTGAG CGACGAGGCT GCACTATCAA CAATAGCCTA 720
TCCACATCCG AGTGGCTTCG AAGATGAGT TCTGATGGAC TGGCTCCCG CAGCATCAAG 780
CAAGAGATGG AGGAAAGGA GAATTGTAC CTGGAGCAGC GACAGGTTAA GGTGAGGAG 840
CCTTCGAGAC CATCAGCTGC CTGGCAGAAC TCTGTGCTGT AGCGGCCACC TACTCTTAC 900
ATGCCCATGA TACAATTCCG CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
ATCTATAGCT GGATGAGGA CCACCTTCCC TACTTTAAGC ACATTGCCAA GCCAGGCTGG 1020
AAGAACTCCG TCGCCACAAA CCTTCCCTG CACGACATGT TTGTCGGGA GACGCTGACC 1080
AATGGCAAGG TCTCCTTCTG GACCATTAC CCCAGTGCCA ACGCTACTT GACATTGGAC 1140
CAGGTGTTTA AGCCACTGGA CCCAGGCTCT CCACAAATGC CCGAGCACTT GGAATCACAG 1200
CAGAAACGAG CGAATCCAGA GCTCCGCGCG AACATGACCA TCAAAACGGA ACTCCCTG 1260
GGCGCACGGG GGAAGATGAA GCCACTGCTA CCACGGGTCA GCTCATACCT GGTACCTATC 1320
CAGTTCCTCG TGAACAGTCT ACTGGTGTG CAGCCCTCGG TGAAGGTGCC ATTGCCCTG 1380
CGCGCTTCCC TCATGAGCTC AGAGCTTGCC CGCCATAGCA AGCGAGTCCG CATTGCCCCC 1440
AAGGTTTTTG GGGACAGGT GGTGTTTGGT TACATGAGTA AGTTCCTTAG TGGCGATCTG 1500
CGAGATTTTG GTACACCCAT CACCAGCTTG TTTAATTTTA TCTTCTTTG TTTATCAGTG 1560
CTGCTAGCTG AGGAGGGAT AGCTCTCTT TCTTCTGAG GACCAGGGAA AGAGGAGAAA 1620
CTCCTGTTTG GAGAAGGGTT TTCTCCTTG CTTCAGTTC AGACTATCAA GGAGGAAGAA 1680
ATCCAGCCTG GGGAGGAAAT GCCACACTTA GCGAGACCCA TCAAGTGGA GAGCCCTCCC 1740
TTGGAAAGAT GGCCTTCCCC GGCCTCATCT TTCAAGAGAG AATCATCTCA CTCTCGGAG 1800
GATTCGTCCC AATCTCCAC CCAAGACCC AAGAAGTCTT ACAGTGGCT TAGTCCCCA 1860
ACCGGTGTG TCTCGGAAT GCTGTGATT CAACACAGGG AGAGGAGGGA GAGGAGCCGG 1920
TCTCGGAGGA AACAGCATCT ACTGCTTCCC TGTGTGATG AGCCGAGCT GCTCTTCTCA 1980
GAGGGGCCCC GTACTTCCC GTCGGCCGCA GAGCTCCCGT TCCAGCAGA CTCTCTGAC 2040
CCTGCTTCCC AGCTCAGTCA CTCCCAGGAA GTGGGAGGAC CTTTAAAGC ACCCATTAAG 2100
GAAAGCTGCG CCATCTCTCT CACCCGAGC AARTCTGTCC TCCCAGAAC CCTGAAATCC 2160
TGGAGGCTCA CGCCCCAGC CAAAGTAGGG GGAATGGATT TCAGCCCATG ACAAACCTCC 2220
CAGGGTGCTT CTGACCCCTT GCCTGACCCC CTGGGGCTGA TGGATCTCAG CACCACTCCC 2280
TTGCAAAAGT TCCCCCTCTG TGAATCACCG CAAAGGCTCC TCAGTTCAGA ACCCTTAGAC 2340
CTCATCTCCG TCCCTTTTGG CAACTCTTCT CCTTCAGATA TAGACGTCCC CAAGCCAGGC 2400
TCCCGGAGC CACAGGTTTC TGGCCTTGCA GCCAATCGTT CTCTGACAGA AGGCCTGTGT 2460
CTGACACAAA TGAATGACAG CCTCAGCAAG ATCTGTCTGG ACATCAGCTT TCCTGGCCTG 2520
GAGCAGGACC CACTGGGCCC TGACAACATC AACTGTGCTC AGTTTATTCC TGAGCTACAG 2580
TAGAGCCCTG CCTTGGCCCC TGTGCTCAAG CTGTCCACCA TCCGGGGCAC TCCAAGGCTC 2640
AGTGACACCC AAGCCTCTGA GTGAGGACAG CAGGCAGGGA CTGTCTGCT CCTCATAGCT 2700
CCTGCTGCC TGATTATGCA AAAGTAGCAG TCACACCTTA GCCACTGCTG GGACCTTGTG 2760
TTCCCAAGA GTATCTGATT CCTCTGCTGT CCTGCCAGG AGCTGAAGGG TGGGAACAAC 2820
AAAGGCAATG GTGAAAGAG ATTAGGAACC CCCAGCCCTG TTTCCATTCT CTGCCAGCA 2880
GTCTCTTACC TTCCCTGATC TTTGAGGGT GGTCCGTGTA AATAGTATAA ATTCTCCAAA 2940
TTATCCTCTA ATTATAAATG TAAGCTTATT TCCTTAGATC ATTATCCAGA GACTGCCAGA 3000
AGGTGGGTAG GATGACCTGG GGTTCATATT GACTTCTGTT CCTTGCTTTT AGTTTGTATA 3060
GAAGGGAAGA CTGCACTGC ACGGTTTCTT CCAGGCTGAG GTACCTGGAT CTTGGGTCTT 3120
TCACTGACAG GACCCAGACA AGTGATCTG CTGCCAGAG TCCTTTTTCG CCTTCCCTGC 3180
CACTTCCCAG TGTTCCTAAG TCAGCTTTCG TGCAAGAAGA AATCCTGGT AAAAAGTCT 3240
TTTGATTGTT GTCAGGAGTT GAATTTGGGG TGGGAGGATG GATGCAACTG AAGCAGAGTG 3300
TGGGTGCCCA GATGTGCGCT ATTAGATGTT TCTCTGATAA TGTCCCAAT CATACCAGGG 3360
AGACTGGCAT TGACGAGAAC TCAGGTGGAG GCTTGAGAAG GCGAAAGGG CCCCTGACCT 3420
GCTTGGCTTC CTAGCTTGC CCTCAGCTT TGCAAGAGC CACCTAGGC CCCAGCTGAC 3480
CGCATGGGTG TGAGCCAGCT TGAGAACACT AACTACTCAA TAAAGCGAA GGTGGACAAA 3540
AAAAA AAAA

```

Seq ID NO: 73 Protein sequence:
Protein Accession #: AAC51128.1

1 11 21 31 41 51
5 MKTSPPRLPI LKRRRLPLPV QNAPSETSEE EPKRSPAQQE SNQAEASKEV AESNSCKFPA 60
GIKIIINHEPTM PNTQVVAIPN NANIHSIITA LTAKGKESGS SGPKNKFIIS CGGAPTQPPG 120
LRPQTQTSYD AKRTEVTLET LGPKPAARDV NLPRPPGALC EQKRETCADG EAAGCTINNS 180
LSNIQWLRIM SSQGLGSRSI KQEMEKEKENC HLEQRQVKVE EPSRPSASWQ NSVSRPPYS 240
YMAIOPAIN STERKMTLK DIYTWIEDHF PYFKHIAKPG WKNRIRHNL LHMDFVRETS 300
10 ANGKVSFWTI HPSANRYLTL DQVFKPLDPG SPQLPEHLES QOKRPNPELR RNMTIKTELP 360
LGARRRMKPL LPRVSSYLVP IQPPVNSQLV LQPSVKVPLP LAASLMSSSEL ARHSKRVRIA 420
PKVFGQVVF GYMSKFFSGD LRDFGTPTS LFNFIPLCLS VLLAEGLAP LSSAGPGKEE 480
KLLFGEQFSP LLPVQTIKEE EIQFGEEMPH LARPIKVESF PLEEWSPAP SFKESSHWS 540
EDSSQSPTFR PKKYSGLRS PTRCVSEMLV IOHRRERRS RSRKQHLPL PCVDEPELLF 600
15 SEGPFSTRWA AELFPFADSS DPASQLSYSQ EVGGPFKTP I KETLPISSST SKSVLPRTPE 660
SWRLTPPAKV GGLDFSPVQT SQGASDPLD PLGLMDLSTT PLQSAPPLES PQRLLSSEPL 720
DLISVFPFNS SPSPIDVPIK GSPEPQVSGL AANRSLTEGL VLDTMNDSL KILLDISPPG 780
LDEDPLGPDN INWSQPIPEL Q

20 Seq ID NO: 74 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 111-416

1 11 21 31 41 51
25 GGAAGAGCC AGGCTGAGCC TTATAAGGA CTGCTCTTTG TCCAAACACA CACATCTCAC 60
TCATCCTTCT ACTCGTGACG CTTCGCCAGCT CTGGCTTTT GAAAGCAAAG ATGAGCAACA 120
CTCAAGCTGA GAGGTCCATA ATAGGCATGA TCGACATGTT TCACAAATAC ACCGACGCTG 180
ATGACAAGAT TGAGAACCCA AGCCTGCTGA CGATGATGAA GGAGAACTTC CCCAATCTCC 240
TTAGTGCCCTG TGACAAAAG GGCACAAAT ACCTCGCCGA TGTCTTTGAG AAAAAGGACA 300
30 AGAATGAGGA TAAGAAGATT GATTTTCTG AGTTTCTGTC CTGCTGGGA GACATAGCCA 360
CAGACTACCA CAAGCAGAGC CATGGAGCAG CGCCCTGTT CCGGGGCAGC CAGTGACCCA 420
GCCCCACCAA TGGGCCTCCA GAGACCCAG GAACAATAA ATGTCTTCTC CCACCAGA

35 Seq ID NO: 75 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
40 MSNTQAERSI IGMDMFHYK TRDDKIEKP SLLTMMKENF PNFLSACDKK GTNYLADVFE 60
KKDKNEDKKI DFSEFLSLLG DIATDYHKQS HGAAPCSGGS Q

45 Seq ID NO: 76 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 111-416

1 11 21 31 41 51
50 GGAAGAGCC AGGCTGAGCC TTATAAGGA CTGCTCTTTG TCCAAACACA CACATCTCAC 60
TCATCCTTCT ACTCGTGACA CTTCGCCAGT CTGGCTTTT GAAAGCAAAG ATGAGCAACA 120
CTCAAGCTGA GAGGTCCATA ATAGGCATGA TCGACATGTT TCACAAATAC ACCGACGCTG 180
ATGACAAGAT TGAGAACCCA AGCCTGCTGA CGATGATGAA GGAGAACTTC CCCAATCTCC 240
TCAGTGCCCTG TGACAAAAG GGCATACATT ACCTCGCCAC TGTCTTTGAG AAAAAGGACA 300
AGAATGAGGA TAAGAAGATT GATTTTCTG AGTTTCTGTC CTGCTGGGA GACATAGCCG 360
55 CAGACTACCA CAAGCAGAGC CATGGAGCG CGCCCTGTT TGGGGGAAGC CAGTGATCCA 420
GCCCCACCAA GGGGCCTCCA GAGACCCAG GAACAATAA TGTCTCTTCC CACCAGA

60 Seq ID NO: 77 Protein sequence:
Protein Accession #: XP_048124.1

1 11 21 31 41 51
65 MSNTQAERSI IGMDMFHYK TGRDGKIEKP SLLTMMKENF PNFLSACDKK GIHYLATVFE 60
KKDKNEDKKI DFSEFLSLLG DIAADYHKQS HGAAPCSGGS Q

70 Seq ID NO: 78 DNA sequence
Nucleic Acid Accession #: Z73678.1
Coding sequence: 253-2433

1 11 21 31 41 51
75 GGGTGGTGC AGGGCAGGGG TGGTATATCC TGTCTGACGG AGGGCGGGCC TGGCCAGTGC 60
CAGAGAGGGA CGAACCAGGG TGGAGAGGCC AGGAGCAGCT GCAGGAGGCC CTCACGCGGA 120
CCTGCACTC TATGGCCGTA GGGAGCGGCT GAGAGCGAGA AGAGCAAGCT CCTGCCGCGC 180
CGCTGCACCG CACCTCGCCT CGCCTCTCTG CTCTCTAGG CCGCGCGCGC GCGCCACCGG 240
CCTCCGCGCA CCATGAACCA CTCGCGGCTC AAGACCGCCT TGGCGTACGA ATGCTTCCAG 300
GACCAGGACA ACTCCAGCTT GGCTTTGCCG TCGGACCAAA AGATGAAAAC AGGCAGCTCT 360
80 GGCAGGCAAG CGGTGACAGA GCAGGTGATG ATGACCGTCA AGCGGCAGAA GTCCAACTCT 420
TCCAGCTCGT CCACCTGAG CCACTCCAAT CGAGGTCCA TGTATGATGG CTGGCTGAC 480
AATTACACT ATGGACCAC CAGCAGGAGC AGCTACTACT CCAAGTTCCA GGCAGGGAAT 540
GGCTCATGGG GATATCCGAT CTACAATGGA ACCCTCAAGC GGGAGCCTGA CAACAGGCGC 600
TTAGCTCTCT ACAGCCAGAT GGAGAACTGG AGCCGCGCAT ACCCCCGGCG CAGCTGTAAC 660
ACCAACCGCG CAGGCAGCA CATCTGCTTC ATGCAGAAAA TCAAGGCGAG CCGCAGTGAG 720
85 CCGAGCTCT ACTGTACCC ACGGGCAACC CTGCGCAAGG GCAOGCTGG CAGCAAGGGC 780
CAGAAGACCA CCCAGAACC CTACAGCTTT TACAGCACT GCAGTGGTCA GAAGGCCATA 840
AAGAAGTGCC CTGTGCGCCC GCCCTCTTGT GCCTCAAGC AGGACCTGT GTATATCCCG 900

	CCCATCTCCT	GCAACAAGGA	CCTGTCTCTT	GGCCACTCTA	GGGCCAGCTC	CAAGATCTGC	960
	AGTGAGGACA	TCGAGTGCAG	TGGGCTGACC	ATCCCCAAGG	CTGTGACAGT	CCTGAGCTCC	1020
	CAGGATGAGA	AGTACCAGGC	CATTGGGGCC	TATTACATCC	AGCATACCTG	CTTCCAGGAT	1080
5	GAATCTGCCA	AGCAACAGGT	CTATCAGCTG	GGAGGCATCT	GCAAGCTGGT	GGACCTCCTC	1140
	CGCAGCCCCA	ACCAGAACGT	CCAGCAGGCC	GCCGCAGGGG	CCCTGCGCAA	CCTGGTGTTC	1200
	AGGAGCAGCA	CCACCAAGCT	GGAGACCCGG	AGGCAGAAAT	GGATCCGCGA	GGCAGTCAGC	1260
	CTCCTGAGGA	GAACCGGGAA	GCOCGAGATC	CAGAAGCAGC	TGACTGGGCT	GCTCTGGAAC	1320
	CTGTCTTCCA	CTGACGAGCT	GAAGGAGGAA	CTCATGCGCG	ACGCCCTGCC	TGTTCTGGCC	1380
	GACCCGCTCA	TCATTCCCTT	CTCTGGCTGG	TGCGATGGCA	ATAGCAACAT	GTCCCGGGAA	1440
10	GTGGTGGACC	CTGAGGTCTT	CTTCAATGCC	ACAGGCTGCT	TGAGGAACCT	GAGCTCGGCC	1500
	GATGCAAGCC	GCCAGACCAT	CGGTAACCTA	TCAGGGCTCA	TTGATTCCCT	CATGGCCTAT	1560
	GTCCAGAACT	GTGTAGCGGC	CAGCCGCTGT	GACGACAAGT	CTGTGGAAAA	CTGCATGTGT	1620
	GTCTCTGACA	ACCTCTCCTA	CCGCTGAGAC	GCGCAGGTGC	CCACCGGCTA	CCGCGAGCTG	1680
	GAGTATAACG	CCCGCAACGC	CTACACCGAG	AAGTCTCTCA	CTGGCTGCTT	CAGCAACAAG	1740
15	AGGCAACAAG	TGATGAACAA	CAACTATGAC	TGCCCCCTGC	CTGAGGAAGA	GACCAACCCC	1800
	AAGGGCAGCG	CGTGGTTGTA	CCATTAGATG	GCCATCCGCA	CCTACCTGAA	CCTCATGGGC	1860
	AAGAGCAAGA	AAGATGCTAC	CCTGGAGGCC	TGTGCTGGTG	CCCTGCAGAA	CCTGACAGCC	1920
	AGCAAGGGCG	TGATGTCCAG	TGGCATGAGC	CAGTTGATTG	GGCTGAAGGA	AAAGGGCCTG	1980
	CCACAAATTG	CCCGCTCTCT	GCAATCTGGC	AACCTCTGAT	TGGTGGGCTC	CGGAGCCTCC	2040
20	CTCCTGAGCC	ACATGTCCCG	CCACCTCTG	CTGCACAGAG	TGATGGGGAA	CCAGGTGTTC	2100
	CCGGAGGTGA	CCAGGCTCCT	CACCAGCCAC	ACTGGCAATA	CCAGCAACTC	CGAAGACATC	2160
	TTGTCTCTCG	CCTGCTACAC	TGTGAGGAAC	CTGATGGCCT	CGCAGCCACA	ACTGGCCCAAG	2220
	CAGTACTTCT	CCAGCAGCAT	GCTCAACAAC	ATCATCAACC	TGTGCCGAAG	CAGTGCCTCA	2280
25	CCCAAGGGCG	CAGAAGCTGC	CCGGCTTCTC	CTGTCTGACA	TGTGGTCCAG	CAAGGAACTG	2340
	CAGGGTGTCC	TCAGACAGCA	AGGTTTCGAT	AGGAACATGC	TGGGAACCTT	AGCTGGGGCC	2400
	AACAGCTACA	GGAACTTCA	CTCCGATTTC	TAAGAAGAGA	CTGTCCAAGC	AAGTTAGGCT	2460
	TGCAAGGAAG	TATGACCCAG	CTGAGAAGCC	CTCAGGCCCT	GCTGGATGGG	GTCTTCTGTC	2520
	CATCCTGTGC	AGTATTGTGG	AAAGTTCACA	AGAAACTGAG	AAGAAACCTA	AAAACCTGTG	2580
30	ATAGTGGAAA	GATTTTGTGA	TTTTTTTTTT	CCCTGGGGAA	ACTGGCAGGC	AATGGGGGTT	2640
	AGGGAGTTTG	GGGCGGGGGG	GGCTTCTCTG	AGTTAAAGGG	GCTTATATGT	GATGTCAATA	2700
	TTTCTCTCTC	TGAGAAATGG	TATATATATG	TGCTTAATGT	AAGTGTGTGC	ATGCATGTGC	2760
	GCGTGCATGT	GTGTGTGTGT	GAGTGTCTTA	AAGCATAACC	ACAACTGCA	AAAAGCTAGG	2820
	TAAGCTATTT	TGTTGCAGCT	CATAAGGTGG	TGAAAAGGAC	TCTCTGTGT	TTCTTACTCA	2880
35	TAGCAAGAGA	CACATGTGTC	TTTTTGTGTA	GCTGCTCATA	ATTCTGAAA	TGTTGTGTGC	2940
	CAGGGCAAGG	GGGCACTCAC	TGCAGTCAGG	CCCTCAGAGG	AGTCTCTCAG	GCTTCTTACC	3000
	AGTGTCTCC	AAGGGTCAG	GAGTAACCTG	GGCTGGGCCA	GCCTCCCCCC	TTACAAGGCT	3060
	GCTTTCCAGC	AAGGGAGTTC	TGGTGTATCT	CATGGGAGAA	TCTGGGGTGT	CTGTAGTGTG	3120
	ACCCCTCCAG	CAGCGCCACA	AGGACTGAGG	TTGGGTAGGT	GTGAGGTTC	AGAGGACAGC	3180
40	AGGACACTCT	CGCATACTTT	GCCAAATGAG	GCCTGCTCAG	AGGAGTAGGA	GCTGAAAGAT	3240
	GGTGCCTTCC	ACCTCTCTGG	GCTGTGTGCC	CATCAGAGCA	GGCTCAGCCT	GCAAAGGGCC	3300
	TGCATTGAGA	GGTCTGTGTA	TCTACTTGT	GCAGGAGAAA	GAAGGTAAAA	AATGATTTTT	3360
	TTAAGAAAA	CTATTTTAT	GCAGCTCTTT	CCCAAGAGCT	GTCTGCGGAA	TGGCTGTGCT	3420
	TCATATTCCC	AGTGGAGAGG	GGAAACAAGT	GGGCTGGGCA	TATACCTATT	CCGGCTTCTA	3480
45	GTGGGATGGA	GTGGGGGTAT	AGAAATTAAC	CAGGAAGATG	TTTCCACCAA	GCCTGCTGTG	3540
	AGTCAATTGA	GGGAGTGTCT	GGGTCCACAG	AGACTTGGAC	GGGGGGAGTT	TGGGTAGACT	3600
	AGGAAGGAA	AGTGCCATAT	CAGGGTACCG	GTACCGGCAA	GCTCATATCT	CAGCCAGGGG	3660
	CCATGCCCCA	CTTCCCTGTA	CCCCAGCTGT	CTTGTCTCCA	CTCTGTGAAA	CCCACAGGGG	3720
	ATGTGATAAA	CAGGGCTATT	AGGGGTATCA	GCCACGTCGA	GCCCCCAGAC	TCTGTGCACT	3780
50	TCAGACCCAG	AGCAGCAGGA	GGGCTCCCGA	GGGCCTTATG	AGAAAACTG	TGTGGACATC	3840
	CCTTGGTGTG	CATTAAGACA	GAGCAGAGCC	CAGGCTCCCG	AAGCCTTCTC	CCTTCCAGCT	3900
	TCTACCTCCA	TGCTAGCAT	GCTGGTGTGA	GAGAGGAATT	AACCTCTGCG	TCTGTGCCCT	3960
	TCTCTAGAAG	AATATAAGAT	GCTCCTCTCT	CTCACCCTTT	CTCAGCCTCC	TCCCAAGTCT	4020
	TCCTCTTCTG	CACCACCCCG	GAGTCCAAAC	CCACCTCTTG	CCCCAGCATT	CAGGCTGGAA	4080
55	AACACTGATG	TGGACTCAGT	ATGACAACTG	AGATGGGGGA	AGCCAGACAT	GTGAGGACGC	4140
	TGTCCTCCGA	GAGGTGTCTC	CGGCTGTTAG	CCAGCTGTGC	TGTGGTCTGT	TGGGTCTGTC	4200
	ATACCTCTCC	TGCTCTCTGT	TCACACTGGG	AGGCCCACTC	CTGGCTCACC	TCTCCTCTCT	4260
	AGGGACCCAC	GTGGAGCCCT	GGATCCCTGG	ACTGTCTCTG	GCATAGGTTT	CAGGGGCTCT	4320
60	CTTTGTGTCT	ATCAGAAACCT	AGAGGAATTC	TCTCTCTAAA	AAATACGTAT	GGCATACCAA	4380
	TCTGTGCGGG	CGAGTGTCTT	AAGCACTTAG	ACTACATCAG	GGAGAACAC	AGACCAATC	4440
	CCGCTCTCCA	TGCGGCTTAT	GTCTTCTGGA	GGAAAGTGGG	GACACAAGTC	CTTGGCTTTA	4500
	GGGCTCCCCC	GGCTGGGGGC	TGTGCAGTCC	GGTCAGGGCG	GGAGGGGAAA	TGCACCGCTG	4560
	CATGTGAACC	TTACAGAGCC	AGGCGGATGC	CCCTTCCCTT	TAGCACTACC	CTGGCCTCCT	4620
	GCATCCCTCT	GCCTCATGTT	CCTCCACCT	TCAAAGAAAT	AAGAGCCCCA	TGGGCCAGC	4680
65	CCCTGCGCTG	GGAAACAGGC	AGCCTTCCAG	ACCTCAGGGG	CTGAGGCAGA	CTATTAGGGC	4740
	AGGGCTGACT	TTGGTGACAC	TGCCCAITCC	CTCTCAGGCC	AGCTCAGGTC	ACCCGGGCTT	4800
	CTGACCCAGG	CCTGTCACTT	TGAGAGGGGC	AAAACCTAGA	GGGGCTTTTC	CTAGAGAAAG	4860
	AGAACAAGGA	GCTTGCCAGG	CTTCATGTAG	CCGACACAGC	TCTCAGGATT	TTAAGTCCAC	4920
	ATTGGCCTCA	CATAGCCCTA	GGCCAATGCC	CAAAATAAGG	AGTTCCAATT	TGGGGCCAAA	4980
70	TGAGGAAGGA	CACAGACTCT	GCCCTGGGAT	CTCCTGTGCT	AGGGGCCAAT	GACAAATCCA	5040
	GTCAATTGGC	ACCAGCCACC	TCTGCAGTGG	GGACCAACT	AGCAGCCCTG	ACTCCCACT	5100
	CCTCCTGGGG	ACCCAAGAGG	CAGTGTGTCT	GTCTGCTGT	CCACCTTGGG	ATCTGGCTGA	5160
	ACTGGCTGGG	AGGACCAAGA	CTGGGGCTGG	GSTGGGCAGG	GAAGGGAAGC	CGGGGGCTGC	5220
	TGTAGGGGAT	CTTGGAGCTT	CCCTGTAGCC	CACCTTCCCC	TTGCTTCATG	TTTGTAGAGG	5280
75	AACCTTGTGC	CGGCCAGGCC	CAGTTTCTCT	GTGTGATACA	CTAATGTATT	TGCTTTTTTT	5340
	GGAAATAGAG	AAATCAATA	AATTGCTAGT	GTTTCTTTGA	AAAAAATA		

Seq ID NO: 79 Protein sequence:
Protein Accession #: CAA98022.1

80	1	11	21	31	41	51	
	MNHSPLKAL	AYECPQDQD	STLALPSDQK	MKTGTSGRQR	VQEQVMIVK	RQSKSSQSS	60
	TLSSNRGSM	YDGLADNYN	GTTSRSSYY	KPQAGNGSWG	YPIYNGTLKR	EPDNRFPSSY	120
85	SQMNWSRHY	PRGSCNTTGA	GSDICFMQKI	KASRSEPDLY	CDPRGTLRKG	TLGSKGQKTT	180
	QNRYSFYSTC	SGQKAIKCP	VRPPSCASKQ	DPVYIPPI9C	NKDL9F9HSR	ASSKICSEDI	240
	EC9GLTIPKA	VQYLS9QDEK	YQAIGAYYIQ	HTCFQDE9AK	QVYQ9LGGIC	KLVDLLRSPN	300
	QNVQ9AAGA	LRNLVFRSTT	NKLETRRQNG	IREAV9LLRR	TGNA9IQKQL	TGLLWNL9ST	360

DELKEELIAD ALFVLADRVI IPFSGWCDGN SNMSREVVDV EVFFNATGCL RNLSSADAGR 420
QTMRYNSGLI DSLMAYVQNC VAASRCDDKS VENCVCVLHN LSYRLDAEVP TRYRQLEYNA 480
RNAYTEKST GCFSNKSDKM MNMNYDCPLP EETNPKGSG WLYKSDAIRT YLNLMGKSKK 540
DATLEACAGA LQNLTAASKGL MSSGMSQLIG LKEKGLPQIA RLQSGNSDV VRSGASLLSN 600
MSRHPLLRHV MGNQVFPVET RLITSHRTGNT SNSEDILSSA CYTVRNLMAS QPQLAKQYFS 660
SSMLANNIINL CRSSASPKAA EAARLLLSMD WSSKELQGV L RQQGFDRNML GTLAGANSLR 720
NFTSRF

10

Seq ID NO: 80 DNA sequence
Nucleic Acid Accession #: NM_006516.1
Coding sequence: 180-1658

15

1 11 21 31 41 51
TAGTCGCGGG TCCCCGAGTG AGCAAGCCAG GGAGCAGGAG ACCAAACGAC GGGGGTCGGA 60
GTCAGAGTCG CAGTGGGAGT CCCCAGGACG GAGCAGCAGC CTGAGCGGGA GAGCGCGCT 120
CGCAGCGCGG TCGCCACCGG CGTACCCGGC GCAGCCAGAG CCACCAAGCG AGCGCTGCCA 180
TGGAGCCGAG CAGCAAGAAG CTGACGGGTC GCCTCATGCT GGCTGTGGGA GGAGCAGTGC 240
TTGGCTCCCT CGAGTTTGGC TACAACACTG GAGTCATCAA TGCCCCCAG AAGGTGATCG 300
AGGAGTTCTA CAACCAAGACA TGGGTCCACC GCTATGGGGA GAGCATCCTG CCCACCAAGC 360
TCACCACGCT CTGGTCCCTC TCAGTGGCCA TCTTTTCTGT TGGGGGCATG ATTGGCTCCT 420
TCTCTGTGGG CTTTTCGTTT AACCGCTTGT GCCCGCGGAA TTCAATGCTG ATGATGAACC 480
TGCTGGCCTT CGTGTCCGCC GTGCTCATGG GCTTCTCGAA ACTGGGCAAG TCCTTTGAGA 540
TGCTGATCCT GGGCGGCTTC ATCATCGGTG TGTACTGGGG CCGACCCACA GGCTTGTGTC 600
CCATGTATGT GGGTGAAGTG TCACCCACAG CTTTTCGTGG GGCCCTGGGC ACCCTGCACC 660
AGCTGGGATC GTCTGTGGG ATCCTCATCG CCCAGGTGTT CGGCGTGGAC TCCATCATGG 720
GCAACAAGGA CCTGTGGCCC CTGCTGCTGA GCATCATCTT CATCCCGGCC CTGCTGCAGT 780
GCATCGTGCT GCCCTTCTGC CCGGAGAGTC CCGCTTCTCT GCTCATCAAC CGCAACGAGG 840
AGAACCGGGC CAGTGTGGTG CTAAGAAGAG TGCGCGGGAC AGCTGACGTG ACCCATGACC 900
TGCAGGAGAT GAAGGAAGAG AGTCGGCAGA TGATGCGGGA GAAGAAGGTC ACCATCCTGG 960
AGCTGTTCGG CTCCCGCGCC TACCGCCAGC CCACTCTCAT CGCTGTGGTG CTGCAGCTGT 1020
CCCAGCAGCT GTCTGGCATC AACGCTGTCT TCTATTACTC CACGAGCATC TTCGAGAAGG 1080
CGGGGGTGCA GCAGCGTGTG TATGCCACCA TTGGCTCGGG TATGTCACAC ACGGCTTCA 1140
CTGTGCTGTC GCTGTTTGTG GTGGAGCGAG CAGGCGGGCG GACCTGACAC CTCATAGGCC 1200
TGCTGTGGAT GGGGGGTGTG GCCATACTCA TGACCATCGC GCTAGCACTG CTGAGAGCAG 1260
TACCTGTGAT GTTCTATCTG AGCATGTGTT CCACTCTTGG CTTTGTGGCC TTCTTTGAAG 1320
TGGGTCTGTG CCCCATCCCA TGGTTCATCG TGGCTGAAC CTTCAGCCAG GGTCCACGTC 1380
CAGCTGCCAT TGCCGTGTGA GGCTTCTCCA ACTGGACCTC AAATTTTCAT GTGGGCGATG 1440
GCTTCCAGTA TGTGGAGCAA CTGTGTGGTC CCACTCTTCT CATCATCTTC ACTGTGCTCC 1500
TGGTTCGTGT CTTTCATCTC ACCTACTTCA AAGTTCCTGA GACTAAAGGC CGGACCTTCG 1560
ATGAGATGCG TTCCGCGCTC CGGAGGGGGG GAGCCAGCCA AAGTATAAG ACACCGGAGG 1620
AGCTGTGTCA TCCCCGCGG GCTGATTCCC AAGTGTGAGT GCGCCAGAT CACCAGCCCG 1680
GCCGTCTCCC AGCAGCCCTA AGGATCTCTC AGGAGCAGAG GCAGCTGGAT GAGACTTCCA 1740
AACCTGACAG ATGTGAGCGG AGCGGGGCTT GGGGCTCCTT TCTCCAGCCA GCAATGATG 1800
CCAGAAAGAT ATTCAAGGAT TAACGGCTCC AGGATTTTAA CAAAGCAAG ACTGTGCTC 1860
AAATCTATTC AGACAAGCAA CAGGTTTTAT AATTTTCTTA TTAAGTATT TGTATTATTT 1920
ATATCAGCCT GAGTCTCTCT TGCCACATC CCAGGCTTCA CCCTGAATGG TTCCATGCCT 1980
GAGGGTGGAG ACTAAGCCCT GTGAGACAC TTGCTTCTCT CACCCAGCTA ATCTGTAGGG 2040
CTGGACCTAT GTTCAAGGA CACACTAATC GAACATGAA CTACAAGCT TCTATCCAG 2100
GAGGTGGCTA TGCCACCCCG TTCTGTGGC CTGGATCTCC CCACTCTAGG GGTGAGGCTC 2160
CATTAGGATT TGCCCTTCCC CATCTCTTCC TACCCAACCA CTCAAATTAA TCTTCTCTTA 2220
CCTGAGACCA GTTGGGAGCA CTGGAGTGCA GGGAGGAGAG GGGAGGGGCC AGTCTGGGCT 2280
GCCGGGTCTC AGTCTCCTTT GCACTGAGGG CCACACTATT ACCATGAGAA GAGGGCCTGT 2340
GGGAGCCTGC AAATCACTG CTCAAGAAGA CATGGAGACT CCGCCCTGT TGTGTATAGA 2400
TGCAAGATAT TTATATATAT TTTTGGTGT CATATATATA TACAGACACT AAGTATAGT 2460
ATATCTGGAG AAGCCAACTT GTAATACAC CACCTCACTC CTGTACTTCA CCTAAACAGA 2520
TATAAATGGC TGGTTTTATG AAACATGGTT TTGAAATGCT TGTGGATTGA GGGTAGGAGG 2580
TTTGGATGGG AGTGAGACAG AAGTAAGTGG GGTGCAACC ACTGCAACGG CTAGACTTCT 2640
GACTCAGAGT CCACTCCCTT ACACGTACCT CTCATCAGTG TCCCTCTGCT CAAAAATCTG 2700
TTTGATCCCT GTTACCAGCA GAATATATAC ATTCTTATTC TTGACATCA AGGCATTCTT 2760
ATACATATT TGATAGTTGG TGTTCAAAA AACACTAGTT TTGTGCCAGC CGTGATGCTC 2820
AGGCTTGAAG TGCATTATT TTGAATGTGA AGGGAA

65

Seq ID NO: 81 Protein sequence:
Protein Accession #: NP_006507.1

70

1 11 21 31 41 51
MEPSSKILTG RLMLAVGGAV LGSLOFGYNT GVINAPQKVI EEFYNQTVWH RYGESILPTT 60
LITLWLSLVA IPSVGGMIGS FSVGLPVNRF GRRNSMLMNN LLAPVSAVLM GFSKLGKSPF 120
MLILGRFIIG VYCGLTTFV PMYVGEVSP APRGALGTLH QLGIIVGILI AQVFGLD SIM 180
GNKDLWPLLL SIIFIPALLQ CIVLPFCPE PRPLLINRNE ENRAKSVLKK LRGTADVTHD 240
LQEMKEESRQ MMREKKVITL ELFRSPAYRQ PILIAVLQL SQQLSGINAV FYYSTSIFEK 300
AGVQQPVTAT ICGGIUNTAP TVVSLPVVER AGRRTLHLIG LAGMAGCAIL MTIALALLEQ 360
LPWMSYLSIV AIPGFVAPFE VGPPIPWFI VAEPLSQQPR PAALAVAGPS NWTNFIIVGM 420
CPQYVEQLCG PYVFIIFTVL LVLFFIFTYF KVPETKGRTP DEIASGFRQG GASQSDATPE 480
ELFPLGADS QV

80

Seq ID NO: 82 DNA sequence
Nucleic Acid Accession #: BC001291
Coding sequence: 44-541

85

1 11 21 31 41 51
GGGGGCGCGG CGCGCTGACC CTCCTGGGCG ACCGCTGGGG ACGATGGCGC TGCTCGCCTT 60
GCTGCTGGTC GTGGCCCTAC CGCGGGTGTG GACAGACGCC AACCTGACTG CGAGACAACG 120

	AGATCCAGAG	GACTCCGAGC	GAACGGACGA	GGGTGACAAT	AGAGTGTGGT	GTGATGTTTG	180
	TGAGAGAGAA	AACACTTTG	AGTGCCAGAA	CCCAAGGAGG	TGCAAAATGA	CAGAGCCATA	240
	CTGCGTTATA	GCGGCCGTGA	AAATATTTCC	ACGTTTTTTC	ATGGTTGCGA	AGCAGTGCTC	300
5	CGCTGGTTGT	GCAGCGATGG	AGAGACCCAA	GCCAGAGGAG	AAGCGGTTTC	TCCTGGGAAG	360
	GCCCATGCCC	TTCTTTTACC	TCAAAGTGTG	TAAAATTGCG	TACTGCAATT	TAGAGGGGCC	420
	ACCTATCAAC	TCATCAGTGT	TCAAAGAATA	TGCTGGGAGC	ATGGGTGAGA	GCTGTGGTGG	480
	GCTGTGGCTG	GCCATCCTCC	TGCTGCTGGC	CTCCATTGCA	GCCGGCCTCA	GCCTGTCTTG	540
	AGCCACGGGA	CTGCCACAGA	CTGAGCCTTC	CGGAGCATGG	ACTCGCTCCA	GACCGTTGTC	600
10	ACCTGTTTGA	TAAACTTGT	TTTCTGTTGA	TTACTCTCTG	GTTTGACTTC	CCAGGGTCTT	660
	GGGATGGGAG	AGTGGGGATC	AGGTGCAATT	GGCTCTTAAC	CCTCAAGGGT	TCTTTAACTC	720
	ACATTCAAG	GAAGTCCAGA	TCTCCTGAGT	AGTGATTTTG	GTGACAAGTT	TTTCTCTTTG	780
	AAATCAAACC	TTGTAATCA	TTTATTGCTG	ATGGCCACTC	TTTTCTTGA	CTCCCTCTG	840
	CCTCTGAGGG	CTTCAGTATT	GATGGGGAGG	GAGGCCATAA	TACCACTCAT	GGAGAGTATG	900
15	TGCTGAGATG	CTTCCGACCT	TTCAGGTGAC	GCAGGAACAC	TGGGGGAGTC	TGAATGATTG	960
	GGGTGAAGAC	ATCCCTGGAG	TGAAGGACTC	CTCAGCATGG	GGGGCAGTGG	GGCACACGTT	1020
	AGGGTGCACC	CCATTCCAGT	GGTGGAGGCG	CTGTGGATGG	CTGCTTTTCC	TCAACCTTTT	1080
	CTACAGATT	CCAGGAGGCA	GAAGATAACT	AATTGTGTTG	AAGAACTTA	GACTTCACCC	1140
	ACCAGCTGGC	ACAGGTGCAC	AGATTCTATA	ATTCCACAC	GTGTGTGTTT	AACATCTGAA	1200
20	ACTTAGGCCA	AGTAGAGAGC	ATCAGGGTAA	ATGGCGTTCA	TTTCTCTGTT	AAGATGCAGC	1260
	CATCCATGGG	GAGCTGAGAA	ATCAGACTCA	AAGTTCACCC	AAAAACAAAT	ACAGGGGAC	1320
	TTCAAAAGTT	CACGAAAAAA	AAAAAAAATA	AAAAAAAATA	AAAAAAAATA	AAA	

Seq ID NO: 83 Protein sequence:

Protein Accession #: AAH01291

25	1	11	21	31	41	51	
	MALLALLLVV	ALPRVWTDAN	LTARQRDPED	SQRTDEGDNR	VWCHVCEREN	TFECQNRRC	60
30	KWTEPYCVIA	AVKIFRPFM	VAKQCSAGCA	AMERPKPEEK	RFLLEEFMPF	PYLKCKKIRY	120
	CNLEGPFFNS	SVFKFYAGSM	GESCGGLWLA	ILLLLASIAA	GLSL		

Seq ID NO: 84 DNA sequence

Nucleic Acid Accession #: NM_022893.1

Coding sequence: 229-2726

35	1	11	21	31	41	51	
	TTTTTTTTTT	TTTTTTGCTT	AAAAAAAAGC	CATGACGGCT	CTCCCAACAT	TCATCTTCCC	60
40	TGCGCCATCT	TGTGATTATT	TCTAATTAT	TTTGGATGTC	AAAAGGCAC	GATGAAGATA	120
	TTTTCTCTGG	AGTCTCTCTC	TTTCTAACCC	GGCTCTCCCG	ATGTGAACCG	AGCCGTCGTC	180
	CGCCCGCCGC	CGCCCGCCGC	CGCCCGCCGC	CGCCCGCCGC	AGCCCAACAT	GTCTCGCCGC	240
	AAGCAAGGCA	AACCCACAGA	CTTAAGCAAA	CGGGAATTCT	CGCCCGAGCC	TCTTGAAGCC	300
	ATTCTTACAG	ATGATGAACC	AGACCACGGC	CCGTTGGGAG	CTCCAGAAGG	GGATCATGAC	360
45	CTCCTCACTT	CTGGGCAGTG	CCAGATGAAC	TTCCCATTTG	GGGACATTCT	TATTTTATAT	420
	GAGCACAAAC	GGAAACAATG	CAATGGCAGC	CTCTGCTTAG	AAAAAGCTGT	GGATAAGCCA	480
	CCTTCCCTCT	CACCAATGGA	GATGAAAAAA	GCATCCAATC	CCGTGGAGGT	TGGCATCCAG	540
	GTCAAGCCAG	AAGATGACGA	TTGTTTATCA	ACGTCTATCT	GAAGAATTGT	CCCCAAACAG	600
	GAACACATAG	CAGATAAACT	TCTGCACTGG	AGGGGCTCTT	CCTCCCTCTG	TCTGACATAT	660
50	GGAGCTCTAA	TCCCAACGCC	TGGGATGAGT	GCAGAATATG	CCCGCAGGGG	TATTTGTAAA	720
	GATGAGCCCA	GCAAGTACAC	ATGTACAAC	TGCAAAACAG	CATTCAACCA	TGCAATGGTTT	780
	CTCTTGCAAC	ACGCAACAGA	CACATCATGA	TTAAGAATCT	ACTTAGAAAG	CGAACACGGA	840
	AGTCCCTCTG	CCCGCGGGGT	TGGTATCCCT	TCAGGACTAG	GTGCAGAATG	TCTTCTCCAG	900
	CCACTCTCTC	ATGGGATTCA	TATTGACAGC	AATAACCCCT	TAAACCTGCT	AAGAATACCA	960
55	GGATCAGTAT	CGAGAGAGGC	TTCCGGCCTG	GCAGAAGGGC	GCTTTCCACC	CATCTCCCCC	1020
	CTGTTTAGTC	GCAGCCAGAG	ACATCACTTG	GACCCCAACC	GCATAGAGCG	CCTGGGGGGG	1080
	GAAGAAATGG	CCCTGGCCAC	CCATCACCCG	AGTGCCTTTG	ACAGGGTGCT	CGGGTTGAAT	1140
	CCAATGGCTA	TGGAGCCTCC	CGCCATGGAT	TTCTCTAGGA	CACTTAGAGA	GCTGGCAGGG	1200
	AACAGGCTCA	GCCCAACGCT	GTCCCCAGGC	CGGCCAGAGC	CTATGCAAGG	GTTACTGCAA	1260
60	CCATTCCAGC	CAGGTAGCAA	GCGGCCCTTC	CTGGCGAGCG	CCCCCTCTCC	TCCTCTGCAA	1320
	TCGCGCCCTC	CTCCCTCCCA	GCCCCCGGTC	AAGTCCAAGT	CATGCGAGTT	CTGCGGCAAG	1380
	ACGTTCAAA	TTCAAGACAA	CCTGGTGGTG	CACCGCGGCA	GCCACACGGG	CGAGAAGCCC	1440
	TACAAGTGCA	ACCTGTGCGA	CCACGGGTGC	ATCCAGGCCA	GCAAGCTGAA	GCGCCACATG	1500
	AAGAGCCACA	TGCACAATTC	GTCCCCCATG	ACGCTCAAGT	CCGACGACGG	TCTCTCCACC	1560
65	GCCAGCTCCC	CGAAGCCCGG	CACCAAGGAC	TTGGTGGGCA	GCGCCAGCAG	CGCGCTCAAG	1620
	TCCGTGGTGG	CCAAGTTCAA	GAGCGAGAAC	GACCCCAACC	TGATCCCGGA	GAACGGGGAC	1680
	GAGGAGGAAG	AGGAGGACGA	CGAGGAAGAG	GAAGAAGAGG	AGGAAGAGAG	GGAGGAGGAG	1740
	CTGACGGAGA	GCGAGAGGGT	GGACTACGGC	TTCCGGCTGA	GCCTGGAGGC	GCGCGGCCAC	1800
	CACGAGAACA	GCTCGCGGGG	CGCGGTGCTG	GGCGTGGGCG	ACGAGAGCCG	CGCCCTGCCC	1860
70	GACGTCTATC	AGGGCATGGT	GCTCAGCTCC	ATGCAGCACT	TCAGCGAGGC	CTTCCACCAAG	1920
	GTCTTGGGCG	AGAAGCATAA	GCGCGGCCAC	CTGGCCGAGG	CCGAGGGCCA	CAGGGACACT	1980
	TGCGACGAAG	ACTCGGTGGC	CGCGGAGTCG	GACCGCATAG	ACGATGGCAC	TGTTAATGGC	2040
	CGCGGCTGCT	CCCCGGGCGA	GTCCGGCTCG	GGGGGCTCTG	CCAAAAAGCT	GCTGCTGGGC	2100
	AGCCCCAGCT	CGCTGAGCCC	CTTCTCTAAG	CGCATCAAGC	TCGAGAAGGA	GTTGACCTTG	2160
75	CCCCCGGCCA	CGATGCCCAA	CACGGAGAAC	GTGTACTCGC	AGTGGCTCGC	CGGCTACGCG	2220
	GCCTCCAGGC	AGCTCAAAGA	TCCCTTCCTT	AGCTTCGGAG	ACTCCAGACA	ATCGCTTTT	2280
	GCCTCTCTGT	CGGAGCACTC	CTCGGAGAAC	GGGAGCTTGC	GCTTCTCCAC	ACCGCCCGGG	2340
	GAGCTGGAGC	GAGGGATCTC	GGGGCGCAGC	GGCACGGGAA	GTGGAGGGAG	CACGCCCAT	2400
	ATTAGTGGTC	CGGGCACGGG	CAGGCCACGC	TCAAAAGAGG	GCAGACGCGA	CGACACTTGT	2460
80	GAGTACTGTG	GGAAAGTCTT	CAAGAACTGT	AGCAATCTCA	CTGTCCACAG	GAGAAGCCAC	2520
	ACGGGCGAAA	GGCCTTATAA	ATGCGAGCTG	TGCAACTATG	CCTGTGCCCA	GAGTAGCAAG	2580
	CTCACCAGGC	ACATGAAAC	GCATGCCAG	GTGGGGAAGG	ACGTTTACAA	ATGTGAATTT	2640
	TGTAAGATGC	CTTTTAGCGT	GTACAGTACC	CTGGAGAAAC	ACATGAAAAA	ATGGCACAGT	2700
	GATCGAGTGT	TGAATAATGA	TATAAAACT	GAATAGAGGT	ATATTAATAC	CCCTCCCTCA	2760
85	CTCCCACTGT	ACACCCCTCT	TTTCACCACT	CCCTTTCCCT	ATCGCCCTCC	AGCCCACTC	2820
	CCTGTAGGAT	TTTTTTCTAG	TCCCATGTGA	TTTAAACAAA	CAAAACAAAC	AACAGAAATG	2880
	ACGAAGCTAA	GAATATGAGA	GTGCTTGTCA	CCAGCACACC	TGTTTTTTTT	CTTTTCTCTT	2940
	TTCTTTTTTC	TTTTTCTCTT	TTTTTTTTTT	TCCTTTATGT	TCTCACCGTT	TGAATGCATG	3000

	ATCTGTATGG	GGCAATACTA	TTGCAITTTA	GGCAAACTTT	GAGCCTTTCT	CTGTGCAAT	3060
	AAITTTACATG	TTGTGTATGT	TTTTTTTAA	ACTTAGACAG	CATGTATGGT	ATGTTATGGC	3120
	TAITTTTAAAT	TGTCCTTAAT	TCGTGTCTGA	GCAAAACATGT	TGCTGTTCCT	AGTTCGGTTC	3180
5	TGAGAGAAA	AGAGAGAGAG	AGAGAAAAAG	ACCATGCTGC	ATACATTCTG	TAATACATAT	3240
	CAITGTACAGT	TTTATTTTAT	AACGTGAGGA	GGAAAAACAG	TCTTTGGATT	AACCCCTCTAT	3300
	AGACAGATA	GATAGCACTG	AAAAAAATC	TCTATGAGCT	AAATGTCTGT	CTCTAAAGGG	3360
	TTAAATGTAT	CAATTGGTAA	GGAGAAAAAA	AGGCCTTGAA	TTGACAAATT	AACAGAAAAA	3420
	CAGAACAAGT	TTATTCCTATC	ATTGGGTTTT	AAATATAGAG	TGCCTTGGAT	CTATTAAAC	3480
	CACATCGATG	GTTCTTTCTA	CTTGTATATA	ACTTGTAGCT	TAATTCAGCA	TGGGTGAGG	3540
10	TAATAAACCT	TAGGAACCTAG	CATATAATTC	TATATTGTAT	TTCTCACAAC	AATGGCTACC	3600
	TAAAAGATG	ACCCATTATG	TCCTAGTTAA	TCATCATTTT	TCCTTTAGTT	TAATTTTATA	3660
	AACAAAAGT	ATTATACAG	TATAAAAGCT	ACTTTGCTCC	TGGTGAGAGC	TAAAAAGAAA	3720
	TGGGCTGTTT	TGCCCAAGT	TTTATTTTTT	TTAAACAATG	ATTAATTTGA	ATGTGTAATG	3780
15	TGCAAAAGCC	CTGGAACGCA	ATTAATATCA	CTAGTAAGGA	GTTCAITTTA	TGAAGATATT	3840
	TGCTTTAATA	TGCTCTTTTT	AAAAATACTG	GCACCAAAAG	AAATAGATCC	AGATCTACTT	3900
	GTTTGTCAAG	TGGAACAATCA	AATGATAAAC	TTTAAGACCT	TGTATACCAT	ATTGAAAGGA	3960
	AGAGGCTGAC	CAATTAAGTTC	GACAGAGGGG	AACAGAGAA	AATAATATGA	TTTATTAGCA	4020
	CAACGTGATA	CTATTGGCCA	TTTAAACTA	GAACAGGTAT	ATAAGCTAAT	ATTGATACAA	4080
20	TGATGATTAA	CTATGAATTC	TTAAGACTTG	CAITTTAAATG	TGACATTCTT	AAAAAAGAA	4140
	GAGAAAGAA	TACTTAAGGG	GCAGTATATA	TGCTGTGTCT	CCCTAAAAGT	TGTACTTCAT	4200
	TTCTTTTCCA	TACACTGTGT	GCTATTGTG	TTAACATGGA	AGAGGATTCA	TGTGTTTTAT	4260
	TTTTATTTTT	TTAATTTTTT	CTTTTTTAIT	AAGCTAGCAT	CTGCCCCAGT	TGGTGTTCAA	4320
	ATAGCACTTG	ACTCTGCCGT	TGATATCTGT	ATCTTTTCTC	TAATCAGAGA	TACAGAGGTT	4380
25	GAGTATAAAA	TAAACCTGCT	CAGATAGGAC	AAITTAAGTG	ACTGTACAA	TTTCCCAGTT	4440
	TACAGGTCTA	TACTTAAGGG	AAAAGTTGCA	AGAATGCTGA	AAAAAAATTG	AACACAATCT	4500
	CATTGAGGAG	CAITTTTAA	AAACTAAAAA	AAAAAAACT	TTGCCAGCCA	TTTACTTGAC	4560
	TATTGAGCTT	ACTTACTTGG	ACGCAACATT	GCAAGCGCTG	TGAATGGAAA	CAGAATACAC	4620
	TTAACATAGA	AATGAATGAT	TGCTTTGCT	TCTACAGTGC	AAGGATTTTT	TGTACAAAA	4680
30	CTTTTTTAAA	TATAAATGTT	AAGAAAAATT	TTTTTTAAAA	AACACTTCAT	TATGTTTAGG	4740
	GGGAACTGAC	ATTTTAGGTT	TCCATTGTCT	TGGTGGTGT	ACAGAGCTTG	TTATCCATT	4800
	AAAAATGATA	GTGGAATTC	TATGCCCTGG	ATACACACCG	CTCTTCAGGT	TGTAAAAAAA	4860
	AAAAACATCA	ATTGGGGAAA	GGTTTAAGAT	TATATAGTAC	TTAAATATAG	GAAAAATGCAC	4920
	ACTCATGTTG	ATTCTATGTC	TAAATACAT	TTATGTCCTT	TTTTCTGTAT	TTCTAGAAATG	4980
35	GTATTGTAAT	TAAATGTTCA	TCTAGTGTTA	GGCACTATAG	TATTTATATT	GAAGCTTGTA	5040
	TTTTTAAGCT	TGCTTGTGTC	TCCTAAAAAG	TATCAATGTA	CCTTTTTTGG	TAGTGGAAAA	5100
	AAAAAGAGCA	GGCTGCCACA	GTATATTTTT	TTAATTGTCG	AGGATAATAT	AGTGCAAAAT	5160
	ATTGTATGTC	TTCAAAAAAA	AAAAAAGAG	AGAAACAAA	AAGTGTGACA	TTACAGATGA	5220
	GAAGCCATAT	AATGGCGGTT	TGGGGGAGCC	TGCTAGAATG	TCACATGGAT	GGCTGTCTATA	5280
40	GGGGTGTGAT	ATATCCTTTT	TTGTTCCCTT	TTCTGCTGTC	CATAGTGTAT	GCAGTACTGC	5340
	AAGCTAATAA	CGTTGGTTTG	TTATGTAGTG	TGCTTTTGT	CCCTTTCCCT	CTATCACCCCT	5400
	ACATTCCAGC	ACTTTACCTT	CATATGCAGT	AAAAGAAAGA	AAGAAAAAAA	AAGGAAAAAA	5460
	AAAAAAAAC	CAATGTTTCA	CAGTTTTTTT	CATTGCCAAA	AACATAATGG	TGCTTTATAT	5520
	TTAGATTGGA	AAGAATTTCA	TATGCAAAGC	ATATTAAAGA	GAAAGCCCGC	TTTAGTCAAT	5580
45	ACTTTTTTGT	AAATGGCAAT	GCAGAAATAT	TTGTTATTTG	CCTTTTCTAT	TCCTGTAAATG	5640
	AAAGCTGTGT	GTCGTAACCT	GAATTTTAT	CTTTTACTAT	GGGAGTCACT	ATTTATATAT	5700
	GCTTATGTGC	CCTGTTCAAA	ACAGAGGCAC	TTAATTGTAT	CTTTATTTTT	TCTTTGTTTT	5760
	TATTTTTTTT	TTTATTATGA	TGACCAAAAG	TCATTACAAC	CTGGCTTTTT	ATTGTATTTG	5820
	TTCTGTGCTT	TTGTTAAGTT	CTATTGGAAA	AACCACTGTC	TGTGTTTTTT	TGGCAGTTGT	5880
50	CTGCATTAA	CTGTTTCATC	ACCCATTTTG	TCCCTTTATT	GAAAAAATAA	AAAAAATTA	5940

Seq ID NO: 85 Protein sequence:
Protein Accession #: NP_075044.1

55	1	11	21	31	41	51	
	MSRRKQKQKQ	HLKSKREFSPE	PLEAILTDDE	PDHGLPLGAPE	GDHDLITCGQ	QMNFFPLGDI	60
	LIPFIEKRRKQ	CNGSLCLSKA	VDEPPSPSPFI	EMKKASNPVE	VGIQVTPEDD	DCLSTSSRRRI	120
60	CPKQSHIADK	LLEWRGLSSP	RSAGHALIPT	PGMSAEYAPQ	GICKEPSSY	TCITCKQPPT	180
	SAWFLLOHAQ	NTEGLRIYLE	SEHGSPLTPR	VGIPLGLGAE	CPSPQPLHGI	HIADNNPFNL	240
	LRIPGVSRE	ASGLAEGRFP	PTPPLFSPPP	RHLDLPHRIE	RLGAEMALA	THHPSAFDRV	300
	LRLNPMAMEP	PAMDFSRRLR	ELAGNTSSPP	LSPGRPSPMQ	RLQLPPQPGS	KPPFLATPPL	360
	PPLQSAPPSS	QPFVKSKECE	FCGKTFKFQS	NLVVHRRSHT	GEKPYKCNLC	DHACTQASKL	420
65	KRMKTHMHK	SSPMTVKSD	GLSTASSPEP	GTSDLVGSAS	SALKSVVAKF	KSENDPNLIP	480
	ENGDEEEED	DEEEEEE	EEELTESER	VDYGFGLSLE	AARHENSRR	GAUVGVGDES	540
	RALPDVMQMG	VLSSMQHFSE	APHQVLGERE	KRGHLAERAG	HRDTCDEDSV	AGESDRIDDG	600
	TVNRRGCSFG	ESASGGLSKK	LLLGSPSSLS	PFSKRIKLEK	EPDLPPATMP	NTEVYSQWL	660
	AGYAAERQLK	DPFLSPGDSR	QSPFASSEH	SENGSLRFS	TPPGELDGGI	SGRSGTGSGG	720
70	STPHISGPGT	GRPSSKEGRR	SDTCEYCGKV	PKNCSNLVH	RRSHTGERPY	KCELONYACA	780
	QSSKLTRMK	THQVGVKDVY	KCEICKMPPS	VYSTLEKHK	KWESDRVLAN	DIKTE	

Seq ID NO: 86 DNA sequence
Nucleic Acid Accession #: XM_035292.2
Coding sequence: 53-1576

75	1	11	21	31	41	51	
	GCTCGCTGGG	CCGCGGCTCC	CGGGTGTCCT	AGGCCCGGCC	GGTGCGCAGA	GCATGGCGGG	60
80	TGCGGCGCCG	AAGCGCGCGG	CGCTAGCGCG	GCGGCGCGCC	GAGGAGAAGG	AAGAGGCGCG	120
	GGAGAAGATG	CTGGCGCGCA	AGAGCGCGGA	CGGCTCGCGC	CCGCGAGGCG	AGGGCGAGGG	180
	CGTGACCTCG	CAGCGGACCA	TCACGCTGCT	CAACGGCGTG	GCCATCATCG	TGGGGACCAT	240
	TATCGGCTCG	GGCATCTTCG	TGACGCGCCAC	GGGCGTCTC	AAGGAGGCGA	GCTCGCGCGG	300
85	GCTGGCGCTG	TGGTGTGGG	CCGCGTGGCG	CGTCTCTCC	ATCGTGGCGG	CGCTCTGCTA	360
	CGCGGAGCTC	GGCACCACCA	TCTCCAAATC	GGGCGGCGAC	TACGCTTACA	TGCTGGAGGT	420
	CTACGCTCG	CTGCGCGCTC	TCTTCAAGCT	CTGGATCGAG	CTGCTCATCA	TCCGCGCTTC	480
	ATCGCAGTAC	ATCGTGGCCC	TGGTCTTCGC	CACCTACCTG	CTCAAGCCGC	TCTTCCCCAC	540

CTGCCCGGTG CCCGAGGAGG CAGCCAAAGCT CGTGGCCTGC CTCTGCGTGC TGCTGCTCAC 600
 GGCCGTGAAC TGCTACAGCG TGAAGGCCGC CACCCGGGTC CAGGATGCCT TTGCCGCCGC 650
 CAAGCTGCTG GCCCTGGCCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GGAAGGGTGA 720
 TGTGTCCAAT CTAGATCCCA ACTTCTCATT TGAAGGCACC AAACCTGGATG TGGGGAACAT 780
 TGTGCTGGCA TTATACAGCG GCCTCTTTGC CTATGGAGGA TGGAAATTACT TGAATTTCGT 840
 CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCTG GCCATCATCA TCCTCCCTGCC 900
 CATCGTGACG CTGGTGTACG TGCTGACCAA CCTGGCCTAC TTCACCACC TGTCACCGCA 960
 GCAGATGCTG TCGTCCGAGG CGGTGGCCGT GGAATTCGGG AACTATCACC TGGGCGTCAT 1020
 GTCCGTGATC ATCCCGGTCT TGGTGGGCTT GTCTGCTTC GGCTCCGTCA ATGGGTCCCT 1080
 GTTCACATCC TCCAGGCTCT TCTTGTGGG GTCCCGGAA GGCCACCTGC CCTCCATCCT 1140
 CTCCATGATC CACCCACAGC TCTCACCCC CGTGGCGTCC CTGTGTGTTA CGTGTGTGAT 1200
 GACGCTGCTC TACGCTTCT CCAAGSACAT CTTCTCCGTC ATCACTTCT TCAGCTTCTT 1260
 CAACCTGGCT TGGTGGCCC TGGCCATCAT CGGCATGATC TGGCTGCGCC ACAGAAAGCC 1320
 TGAGCTTGAG CGGCCCATCA AGGTGAACCT GGCCCTGCTT GTGTCTTCA TCCTGGCCTG 1380
 CCTCTTCTG ATGCGGCTCT CTTCTGGAA GACACCCGTG GAGTGTGGCA TCGGCTTCAC 1440
 CATCATCTC AGCGGGCTGC CGTCTACTT CTTGGGGTTC TGGTGGAAA ACAAGCCCA 1500
 GTGGCTCTCC CAGGGCATCT TCTCCAGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGTT 1560
 CCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

Seq ID NO: 87 Protein sequence:
 Protein Accession #: KP_035292.2

1 11 21 31 41 51
 MAGAGPKRRA LAAPAAEKE EAREKMLAAR SADGSAPAGE GEGVTLQRNI TLLNGVAIIV 60
 GTIIGSGIFV TPTGVLKEAG SPGLALVWMA ACQVFSIVGA LCTAELGTTI SKSGGDYAYM 120
 LEVYGSAPAF LKLWIELLII RPSSQYIVAL VFATYLLKPL FPTCPVPBEA AKLVACLCLVL 180
 LLTAVNCYSV KAATRVQDAF AAKLLALAL IILLGFVQIG KGDVSNLDPN PSFEGTKLDV 240
 GNIVLALYSY LFAYGGWNYL NFVTEEMINP YRNLPLAIII SLPIVTLVYV LTNLAYFTTL 300
 SEQMLSSSEA VAVDFGNYHL GVMWIIIPVF VGLSCFPGSVN GSLFTSSRLP FVGSREGHLP 360
 SILSMIHQPQ LTPVPSLVFT CVMILLYAFS KDIPSVINFF SFFNWLCLVAL AIIGMIWLKH 420
 RKPELERPIK VNLALPVFFI LACLFLIAVS FWKTPVECGI GPTIILSLGP VYFPGVWVKN 480
 KPKNLLQGIF STTVLCQKLM QVVPQET

Seq ID NO: 88 DNA sequence
 Nucleic Acid Accession #: NM_005268.1
 Coding sequence: 168-989

1 11 21 31 41 51
 TAAAAAGCAA AAGAATTGCG GGCCGGGTGCG ACACGGGCTT CCCCAGAAAC CTTCCCGGCT 60
 TCTGGATATG AAATTCAGC TGCTTGCTGA GTCTATTGCG CGGCTGCTGG GAGCCAGGAG 120
 AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACGCGTGG GTCCACCATG AACTGGAGTA 180
 TCTTTGAGGG ACTCCTGAGT GGGGTCAACA AGTACTCCAC AGCCTTTGGG CGCATCTGGC 240
 TGTCTCTGCT CTTCTCTTC CGCGTGTGCG TGTACTGTGT GACGCGCGAG CGTGTGTGGA 300
 GTGATGACCA CAAGGACTTC GACTGCAATA CTGCGCAGCC CGGCTGCTCC AAGTCTGCT 360
 TTGATGAGTT CTTCCCTGTG TCCCATGTGC GCCTCTGGGC CCTGCAGCTT ATCCTGGTGA 420
 CATGCCCTCC ACTGCTGCTG GTCATGCAGC TGGCCTACCG GGAGGTTTCA GAGAAGAGGC 480
 ACCGAGAAGC CCATGGGGAG AACAGTGGGC GCCTCTACCT GAACCCCGGC AAGAAGCGGG 540
 GTGGGCTCTG GTGGACATAT GTCTGCAGCC TAGTGTTCAA GCGAGCGGTG GACATGCGCT 600
 TTCTCTATGT GTTCCACTCA TTCTACCCCA AATATATCCT CCTCTGTG GTCAAGTGCC 660
 AGCGAGATCC ATGTCCCAAT ATAGTGGACT GCTTCATCTC CAAGCCCTCA GAGAAGAACA 720
 TTTTCACCTT CTTCTATGTT GCCACAGCTG CCATCTGCAT CCGTCTCAAC CTGCTGGAGC 780
 TCATCTACCT GGTGAGCAAG AGATGCCACG AGTGCCTGGC AGCAAGGAAA GCTCAAGCCA 840
 TGTGCACAGG TCATCACCCC CACGGTACCA CCTCTCTCTG CAAACAGAC GACCTCCTTT 900
 CGGGTGACCT CATCTTCTG GGCTCAGACA GTCATCCTCG TCTCTTACCA GACCGCCCCC 960
 GAGACCATGT GAAGAAAACC ATCTTGTGAG GGGCTGCTGG GACTGTGCTG GCAGGTTTGG 1020
 CCTGGATGGG GAGGCTCTAG CATCTCTCAT AGGTGCAACC TGAGAGTGGG GGAGCTAAGC 1080
 CATGAGGTAG GGGCAGGCAA GAGAGAGGAT TCAGACGCTC TGGGAGCCAG TTCCTAGTCC 1140
 TCAACTCCAG CCACTGCCCC CAGCTCGACG GCACTGGGCC AGTTCCTCCCT CTGCTCTGCA 1200
 GCTCGGTTTC CTTTCTAGA ATGGAATAG TGAGGGCCAA TGC

Seq ID NO: 89 Protein sequence:
 Protein Accession #: NP_005259.1

1 11 21 31 41 51
 MNWSIFEGLL SGVNYSTAF GRWLVLVFI FRVLVVLVTA ERVMSDDHKO FDCNTRQPGC 60
 SNVCFDEFFP VSHVRLWALQ LILVTCPSLL VVMHVAYREV QEKHREAHG ENSGRLYLNP 120
 GKRGGLMWT YVCSLVFKAS VDIAFLYVPH SFYPKYILPP VVKCHADPCP NIVDCPISKP 180
 SEKNIFTLFM VATAICILL NLVELIYLV KRCHECLAAR KAQAMCTGHE PHGTTSSCKQ 240
 DDLLSGDLIF LGSDSHPLLL PDRPRDHVKK TIL

Seq ID NO: 90 DNA sequence
 Nucleic Acid Accession #: NM_002391.1
 Coding sequence: 26-457

1 11 21 31 41 51
 CGGGGGAAGC AGCGGGGCA GCGAGATGCA GCACCGAGGC TTCTCTCTCC TCACCTCTCT 60
 CGCCCTGCTG CGCTCACCT CGCGGTGCG CAAAAAGAAA GATAAGGTGA AGAAGGGCGG 120
 CCGGGGAGC GAGTGGCTG AGTGGGCTG GGGGCGCTGC ACCCCAGCA GCAAGGATTG 180
 CGGCGTGGST TTCGCGAGG CACCTGCGG GGGCCAGACC CAGCGCATCC GGTGCAAGGT 240
 GCCCTGCAAC TGAAGAAGG AGTTTGGAGC GCAGTGAAG TACAAGTTTG AGAACTGGGG 300
 TGGCTGTGAT GGGGGCAGC GCACCAAGT CCGCCAGGC ACCCTGAAGA AGGCGCGCTA 360

CAATGCTCAG TGCCAGGAGA CCATCCGCGT CACCAAGCCC TGCACCCCA AGACCAAAGC 420
AAAGGCCAAA GCCAAGAAAG GGAAGGGAAG GGAAGGGAAG GGAAGGGAAG GGAAGGGAAG 480
GCCCTGCTGT TCAATGAGGG CCTGCGCCAG CCCTCCCTCT CCCAGGCCCC AGATGTGACC 540
CACCAGTGCC TCTGTCTGCG TCGTTAGCTT TAATCAATCA TGCCCTGCGT TGTCCCTCTC 600
ACTCCCCAGC CCCACCCCTA AGTGCCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT 660
TGAGCCTCCC CCAAGCAAT GTGAGTCCCA GAGCCCGCTT TTGTTCTTCC CCACAAATCC 720
ATTACTAAGA AACACATCAA ATAACTGAC TTTTCCCCC CAATAAAGC TCTTCTTTT 780
TAATAT

10 Seq ID NO: 91 Protein sequence:
Protein Accession #: NP_002382.1

1 11 21 31 41 51
15 MQHRGFLLLT LLALLALTSA VAKKDKVKK GPGSECAEW AWGPCTPSSK DCGVGFREGT 60
CGAQTQRIRC RVPCKWKKEF GADCKYKFEN WGACDGGTGT KVRQGLTKKA RYNAQCQETI 120
RVTKPCTPKT KAKAKAKGK GKD

20 Seq ID NO: 92 DNA sequence
Nucleic Acid Accession #: NM_005130.1
Coding sequence: 98-802

1 11 21 31 41 51
25 CTCTACCTGA CACAGCTGCA GCCTGCAATT CACTCCCACT GCCTGGGATT GCCTGGGATT 60
CGTGTGCTCA GAACAGGTTG AACGCCAGC TGACGCCATG AAGATCTGTA GCCTCACCCCT 120
GCTCTCTCTC CTCTACTGCG CTGCTCAGGT GCTCCTGCTG GAGGGGAAAA AAAAAGTGAA 180
GAATGGAGCTT CACAGCAAGG TGGTCTCAGA ACACAAAGGAC ACTCTGGGCA ACACCCAGAT 240
TAAGCAGAAA AGCAGGCCCG GGAACAAAGG CAAGTTTGTG ACCAAGAGCC AAGCCAACTG 300
CAGATGGGCT GCTACTGAGC AGGAGGAGGG CATCTCTCTC AAGGTTGAGT GCCTCAATT 360
GGACCATGAA TTTTCTCTGT TCTTTGCTGG CAATCCAACC TCATGCTTAA AGCTCAAGGA 420
TGAGAGAGTC TATTGGAAC AAGTTGCCCG GAATCTGCGC TCACAGAAAG ACATCTGTAG 480
ATATTTCCAG ACAGCTGTGA AAACCCAGAT GTGCAGAAAG GATTTTCCAG AATCCAGTCT 540
TAAGCTAGTC AGCTCCACTC TATTGGGAA CACAAAGCCC AGGAAGGAGA AAACAGAGAT 600
GTCCCCCAGG GAGCAGATCA AGGCAAAAGA GACCACCCCC TCTAGCCTAG CAGTGACCCA 660
GACCATGGCC ACCAAGCTC CCGAGTGTGT GGAGGACCCA GATATGGCAA ACCAGAGGAA 720
GACTGCCCTG GAGTTCTGTG GAGAGACTTG GAGCTCTCTC TGCACATTCT TCCTCAGCAT 780
AGTGCAAGAG ACCTCATGCT AATGAGGTCA AAAGAGAAAG GGTTCCTTTA AGAGATGTCA 840
TGTCGTAAAT CCTCTGTAT ACTTTAAAGC TCTCTACAGT CCCCCCAAAA TATGAACCTT 900
TGTCGTAAAT GAGTGCAAGC AAATATTAA ACAAGTTTTC TATTTTTCG TTTTGTGTTT 960
TGGAATTTGC CTATTTTTC TTGGATGCGA TGTTCAGAGG CTGTTTCTCG CAGCATGTAT 1020
TTCCATGGCC CACACAGCTA TGTGTTGAG CAGCGAAGAG TCTTTGAGCT GAATGAGCCA 1080
GAGTGATAAT TTCAGTGCAA CGAACTTTCT GCTGAATTAA TGTGAATAAA ACTCTGGGTG 1140
TTTTTCAAAA AAAAAAAAAA AAA

Seq ID NO: 93 Protein sequence:
Protein Accession #: NP_005121.1

1 11 21 31 41 51
50 MKICSLTLLS FLLAAQVLL VEGKKVKVNG LHSKVSEQK DTLGNTQIKQ KSRPGNKGKF 60
VTEDQANCRW AATQEGEGIS LKVECTQLDH EFSCVFAGNP TSCLKLQDER VYWKQVARNL 120
RSQKIDICRY S KTAVKTRVCR KDFPESLKL VSSTLFGNTK PRKEKTEMSF REHIKKGKETT 180
55 PSSLAVTQTM ATKAPECEVD PDMANQRKTA LEFCGETWSS LCTFFLSIVQ DTSC

Seq ID NO: 94 DNA sequence
Nucleic Acid Accession #: NM_012101
Coding sequence: 125-1891

1 11 21 31 41 51
60 CTCTCAGAGC GTGTGCTCTT AGTCTCTGTG GTTGCTGCTC CCACTCCCTG CCGAGAGCGC 60
TGCCAGAAAG GTCACTATTC CTGAACCCCA GCAAGCCTGA AACAGCTCAG CCAAGCACCC 120
TGCGATGGAA GCTGCAGATG CCTCCAGGAG CAACGGGTG AGCCCAAGAG CCAGGGATGC 180
CCGAGAGCCG TCGGGGCCCA GTGGCAGCCT GGAGAATGGC ACCAAGGCTG ACGGCAAGGA 240
TGCCAGAGCC ACCAAGGGCG ACGGCGGGGA GGCAGCTGAG GGCAAGAGCC TGGGCAAGCC 300
CCTGAAGCCA GGGGAAGGTA GGAGCGCCCT GTTGGCGGGC AATGAGTGGC GGCGACCCAT 360
CATCCAGTTT TCGGAGTCCG GGGACGACAA GAACTCCAAC TACTTCAGCA TGGACTCTAT 420
GGAAGGCAAG AGGTGCGCGT ACGCAGGGCT CCACTGCGGG GCTGCCAAGA AGCCACCCGT 480
TACCTTTGCC GAAAGGGGCG ACGTGCAGAA GTCCATTTC TCGAGTCCC GGAAGCCAC 540
GGTGTCCATC ATGGAGCCCG GGGAGACCCG GCGGAACAGC TACCCCGGG CCGACAGGG 600
CCTTTTTC CAAGTCCAAGT CCGCTCCGA GGAGGTGCTG TGCGACTCTT GCATCGGCAA 660
CAAGCAGAGG GCGGTCAAGT CCTGCCTGCT GTGCAGGCC TCCTTCTGCG AGCTGCATCT 720
CAAGCCCCAC CTGGAGGGCG CCGCTTCCG AGACCAACAG CTGCTCGAGC CCATCGGGGA 780
CTTTGAGGCC CGCAAGTGTG CCGTGCATGG CAAGACGATG GAGCTCTTCT GCCAGACCGA 840
CCAGACTCTG ATCTGTATCC TTTGATGTT CCAGGAGCAC AAGAATCATA GCACCGTGAC 900
AGTGAAGGAG GCCAAGGCGG AGAAGGAGAC GGAGCTGTCA CTGCAAAAGG AGCAGCTGCA 960
GCTCAAGATC ATTGAGATTG AGGATGAAGC TGAGAAGTGG CAGAAGGAGA AGGACCCGAT 1020
CAAGAGCTTC ACCACCAATG AGAAGGCCAT CCTGGAGCAG AACTTCCGGG ACCTGGTGG 1080
GGACCTGGAG AAGCAAAAGG AGGAAGTGAG GGCTGCGCTG GAGCAGCGGG AGCAGGATGC 1140
TGTGAGCCAA GTGAAGGTGA TCATGGATGC TCTGGATGAG AGAGCCAAAG TGCTGCATGA 1200
GGACAGCAG ACCCGGAGC AGCTGCATAG CATCAGGAC TCTGTGTTGT TTCTGCAGGA 1260
ATTGAGTGA TTTGATGAGC ATTACTCTCT CCCCCACCC CTGCCCCACT ATCATGTCT 1320
85 GCTGGAGGGG GAGGCGCTGG GACAGTCACT AGGCAACTTC AAGGACGACC TGCTCAATGT 1380
ATGCATGCGC CAGTTGAGA AGATGTGCAA GGCGGACCTG AGCCGTAACT TCATTGAGAG 1440
GAACCATATG GAGAACGGTG GTGACCATCG CTATGTGAAC AACTACACGA ACAGCTTCGG 1500

5
10
15
20
25

```

GGGTGAGTGG AGTGACCCGG ACACCATGAA GAGATACTCC ATGTACCTGA CACCCAAAGG 1560
TGGGGTCCGG ACATCATACC AGCCCTCGTC TCCTGGCCGC TTCACCAAGG AGACCACCCA 1620
GAAGAATTTC AACAACTCTC ATGGCAACAA AGGTAACCTAC ACCTCCCGGG TCTGGGAGTA 1680
CTCCTCCAGC ATTCAGAACT CTGACAATGA CCTGCCCGTC GTCCAAAGGCA GCTCCTCCTT 1740
CTCCCTGAAA GGCTATCCCT CCTCATGCGG GAGCCAAAGC CCAAGGCCCC AGCCCCAGAC 1800
TTGGAATCTT GGCAAGCAGA CTATGCTGTC TCACTACCGG CCATTCTACG TCAACAAGG 1860
CAACGGGATG GGGTCCAAAG AAGCCCATG AGCTCCTGGC GGAAGGAACG AGGCCCCACA 1920
CCCTGCTCTT TCCTCTCTAG CCGTCTGCTC TTGCTTCTA AGCTACTGTG CTGTCTGGG 1980
TGGGAGGGAG CCTGGTCTTG CACTGCCCT CTGCAGCCCT CTGCGAGCCT CTGGGGGCA 2040
GTTCCGGCCT CTCGACTTC CCCACTGGCC AACTTCCATT CAGACTCCTT TCCTGCCTTG 2100
TGACCTCAGA TGCTCACCAT CATTCCTGTG CTCAGAGGCC AACCCATCAC AGGGGTGAGA 2160
TAGGTGGGG CTGCCCCTAA CCGCCAGCC TCCTCCTCTC GGGCTGGATC TGGGGGCTAG 2220
CAGTGAGTAC CCGCATGGTA TCAGCCTGCC TCTCCCGCCC ACGCCTCTGT GTCTCCAGGC 2280
CTATAGACGT TTCCTCCAA GGCCTATCC CCAATGTG TCAGCAGATG CTGGACAGC 2340
ACAGCCACCC ATCTCCCAT CACATGGCCC ACCTCCTGCT TCCAGAGGA CTGGCCCTAC 2400
GTGCTCTCTC TCGTCTTACC TATCAATGCC CAGCATGGCA GAACCTCGAG TGGCCAAGGG 2460
CTGCAGATG AAACCTCTCA GTGCTTGAC ATCACTTAC CCAGGCGGTG GGTCTCCACC 2520
ACAGCCACTT TGAGTCTGTG GTCCCTGGAG GGTGGCTTCT CTGACTGGC AGGATGACCT 2580
TAGCCAAGAT ATTCTCTGT TCCTCTGCT GAGATAAAGA ATTCCCTAA CATGATATAA 2640
TCCACCCATG CAATAGCTA CTGGCCAGC TACCATTTAC CATTGCGCTA CAGAAATTC 2700
TTCAGTCTAC ACTTTGGCAT TCTCTCTGCG GATGGAGTGT GGTGGGCTG ACCGCAAAAG 2760
GTGCTTACA CACTGCCCCC ACCCTCAGCC GTTGGCCCAT CAGAGGCTGC CTCTCTCTT 2820
TGATTACCCC CCATGTTGCA TATCAGGGTG CTCAGGATT GGAGAGGAGA CAAACCAGG 2880
AGCAGCAGC TGGGACATC TCCGCTCTCA ACAGCCCGC GCCTATGGG GCTCTGGAAG 2940
GATGGGCCAG CTTCAGGGG TTGGGAGGG AGACATCCAG CTGGGCTT CCCCTTTGGA 3000
ATAAACCAT GGTCTGTC

```

Seq ID NO: 95 Protein sequence:
Protein Accession #: NP_036233.1

30
35
40
45

```

1 11 21 31 41 51
MEAADASRSN GSSPEARDAR SPSGPGSLE NGTKADGKDA KTTNGHOGEA AEGKSLGSLAL 60
KPGEGRSALF AGNEWRRII QFVESGDDKN SNYFSDMSME GKRSPYAGLQ LGAARKPPVT 120
FAEKGDVRKS IFSESRKPTV SIMPEGETRR NSYPRADTGL FSRKSGSEE VLCDSCIGNK 180
QKAVKSLVCL QASPCRLRLK PHLEGAAPRD HOLLEPIRDF EARKCPVHGK TMELFCQTDQ 240
TCICYLQMFQ EHKNHSTVTV EEAKAEKETE LSLQKEQLQL KIIIEIDEAE KWQKEKDRIK 300
SPTTNEKAIL EQNFRDLVRD LEKQKEEVRA ALBQRBQDAV DQVKVIMDAL DERAKVLHED 360
KQTRQLHSI SDSVLFLEF GALMSNYSLP PPLETYHVL EGEGLGQSLG NFRDILLNVC 420
MRHVEKMKCA DLSRNFIERN HMEGNDHRY VNNYTNPSFG EWSAPDTMRK YSMYLPKGG 480
VRTSYQSPSP GRFTKETQK NFMNLYGTHG NYTSRVWEYS SSIQNSDNDL PVVQGSSEFS 540
LKQYPSLMRS QSPKAPQTM KSGKQTMLSH YRPFYVKNKN GIGSNEAP

```

Seq ID NO: 96 DNA sequence
Nucleic Acid Accession #: NM_080668.1
Coding sequence: 83-841

50
55
60
65
70
75
80
85

```

1 11 21 31 41 51
GGCAGAGGG CAGCGAGTGG CCTTCCCGGT TGGCGCGGCG CCGGGGCGGC GGCCTGGAG 60
GAGCTCGAGA CGAGCCCTAG TTATGTCTGG GAGGCGAAGC CGTCCGAGG GAGCGCTCA 120
GCGCTCCGGG CCAAGGCCCC CATCTCCTAC TAAGCCTCTG CCGAGGTCCC AGCGGAAATC 180
AGGCTCTGAA CTCCCGAGCA TCCTCCCTGA AATCTGGCGG AAGACACCCA GTGCGGCTGC 240
AGTCAGAAAG CCGCTGCTCT TAAAGAGGAT CGTGGCCCAT GCTGTAGAGG TCCACAGCTG 300
CCAATCACCT CGCAGGAGCC CTAGGATTTC CTTTTCTTGG GAGAAAGAAA ACAGACCCCC 360
TGGCAGGGAG CTTACTAAGG AGGACCTTTT CAAGACACAC AGCGTCCCTG CCACCCCCAC 420
CAGCACTCCT CTGCGCAACC CTGAGGCCGA GTCCAGCTCC AAGGAAGGAG AGCTGGACGC 480
CAGAGCACTG GAAATGTCTA AGAAATGTCAG GCGTTCCTAC AGCCGCTGAG AGACCTTGGG 540
CTCTGCTCT ACCTCCACCC CAGGCCGCGG GTCTCTCTTT GGCTTCGAGG GGCTGCTGGG 600
GGCAGAAGAC TTGTCCGAGG TCTGCGCAGT GGTGTGCTCC AAATCACCG AGGTCCCCAG 660
GGTTTGTGCA AAGCCCTGGG CCCAGACAT GACTCTCCCT GGAATCTCCC CACCAACCGA 720
GAAACAGAAA OGTAAAGAGA AGAAAATGCC AGAGATCTTG AAAACGAGC TGGATGAGTG 780
GGCTGCGGCC ATGAATGCCG AGTTTGAAGC TGCTGAGCAG TTTGATCTCC TGGTTGAATG 840
AGATGCAATG GGGGTGTCAC CTGGCCAGAC TCTCCCTCCT GTCCCTGTACA TAGCCACCTC 900
CCTGTGGAGA GGACACTTAG GGTCCCTCC CCGTGTCTTG TTACTGTGT GTGTGCTGGT 960
GCTGCGCATG AGGACTGTCT GCCTTTGAGG GCTTGGGCG CAGCGGCAGC CATCTTGGTT 1020
TTAGGAAATG GGGCGGCTG GCCCAGCCAC TCACTGGTGT CCTGTCTCTT GTCGTCTGT 1080
CCTTCTATC TCCCAAAAGT ACCATAGCCA GTTTCAGAT GGGCCACAGA CTGGGAGGA 1140
GAATCAGTGG CCCAGCCAGA AGTTAAAGGG CTGAGGGTTG AGGTGAGAGG CACCTCTGCT 1200
CTTGTGGGA GGGGTGGCTG CTTGGAATA GGGCCAGGGG CTCTGCCAGC CTGCGCTCT 1260
CCCTCTGAG TTGCTTCTG TTGCTGGCTT TCTTCTTGA CCCACCTGTG TAAAGAGGTT 1320
TTGAGTTCCG TGGGTTTCCC CTTGATTCT GTAAATAGTC CCAGAGAGAA TTCGTGGGCT 1380
GAGGGCAATT CTGTCTTGA GGAAGAAGCT GGAATTCAG CCTGTGAGT CTGATTTG 1440
AAGGATGTAG GGAGCCTTAG TTGGTCTCA GACATAAGT GTGTACTACA CAGAAGCTGT 1500
GTTTCTAGT TCTGGTCTGC TGTGAGATG TTTGTAATAT GCCAGGTTGA TAGGCGCTG 1560
GCTGCTTGA GCAAGGGTG CATTCAGGG TGTGGCCACC AGGTGCTGTG AGTTTCTGTG 1620
GCTCATGGCC TCTGGGCTGG TCCCTTGCAC AGGGCCCAAG CTGAGTCTT ACCACTTGC 1680
TGCAAGGGTG GAAGGTGGCC CCTCTGTCA CCCATACCCA TTTCTTCAA AATAAGTTAC 1740
ACCGAGTCTA CTGGGCCCTA GAAGAGAAAG TTGAAGATC CCAGACCTAC TAGCATTTG 1800
CAACTATGCT GTTAAAGTCC TCGGAAAGTT TCCTCGCGTA CCAGACAGCG GGGGGGCTG 1860
ATAGCAATT TAGTTTGG CCTCCATATC CTCTCAGATG AGAACACTGC CTGATGCAT 1920
CTCATGATCT CTGAGAATT TCCCATCTT TCTCTCTTT CCATCGTGTG GATTCAATAG 1980
TTTGGATTG AAGGTCGCC TGCCCCGAC TCTCTGCGG CACCCCTGGC CATTGTACT 2040
TTTGTGTTT AGAAGTTCGT GGAAGTAGAC GCTGAGGTGT GCAGAGGAGC TGGTGGATAA 2100
CAGAGAAATG CAGGGAAGAT GAGTGCTGG TACGGTACT TGGATGAAC GGTGACGGCC 2160
AGCGGGCCCC TAATAAAACC CTCTGCCAGG TCTGGGAGTC CCAGGCCATC TGCTCAAGC 2220

```


TCTGTGGTTT GTCAGACCTG CAAGCAAGCC CCCTGCTGGG GAAGCCTAGG TGTCCCTGAG 2280
 CTGAACCCGA CTGAAGAACT CTTGTCTCTA CTGGCTGATG CAGCAGAACT CTTGGGAAAT 2340
 GTCTTAGTCC TGCGAATCA GGAGTCACCA GATGATGCAG AGTTGAGATC ATCATTTGCA 2400
 AGTTCTCTGT TCCTGAGGAA CTAAATTTAA GGAAAAAATG GGATTTTGT TTAGAGTTGG 2460
 AAAAAAGCC TGATTAAAGA GTTCTGCTCT GTTAAAAAAA AAAAAA AAAAA

Seq ID NO: 97 Protein sequence:
 Protein Accession #: NP_542399.1

10 1 11 21 31 41 51
 | | | | |
 MSGRRTRSGG AAQRSGPRAP SPTKPLRRSQ RKSGSELPSI LPEIWPKTPS AAARVKPIVL 60
 KRIVAHAVEV PAVQSPRRSP RISFFLEKEN EPPGRELTKE DLFKTHSVPA TPTSTPVPNP 120
 EAESSKSGE LDARDLEMSK KVRRSYSRLT TLGSASTSTP GRRSCFGFEG LLGAEDLSGV 180
 15 SPVVCCKLTE VPRVCAKPWA PDMTLPGISF PPEKQKRKKK RMPEILKTEL DEWAAAMNAE 240
 FEAAEQFDLL VE

Seq ID NO: 98 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 58-12444

20 1 11 21 31 41 51
 | | | | |
 GGGGCATTTC CGGGTCCGGG CCGAGCGGGC GCACGCGCGG GAGCGGGACT CGGCGGCATG 60
 GCGGGCTCCG GAGCCGGTGT GCGTTGCTCC CTGCTGCGGC TGCAGGAGAC CTTGTCCGCT 120
 GCGGACCCCT GCGGTGCTGC CCTGGCCGGT CATCAACTGA TCCGCGGCCT GGGGCAGGAA 180
 TCGCTCTCTG GCGAGCAGCC CGCGTGCTG GCATTACAGA CATCTTTAGT TTTTCCAGA 240
 GATTTGCGTT TGCTTGTATT TGTCGGGAAG TCACTCAACA GTATTGAAT TCGTGAATGT 300
 AGAGAAGAAA TCCTAAAGTT TTTATGTATT TTCTTAGAAA AAATGGGCCA GAAGATGCSA 360
 CTTACTCTG TGGAATTAA GAACACTTGT ACCAGTGTGT ATACAAAAGA TAGAGCTGCT 420
 AAATGTAAAA TTCAGCCCTT GGACCTTCTT ATTAAGTTAC TTCAGACTTT TAGAAGTTCT 480
 AGACTCATGG ATGAATTTAA AATTGGAGAA TTATTAGTA AATTCATGG AGAAGTTGCA 540
 TTGAAAAAAA AAATACCAGA TACAGTTTAA GAAAAAGTAT ATGAGCTCCT AGGATTATTG 600
 GGTGAAGTTC ATCTTAGTGA GATGATAAAT AATGCAGAAA ACCTGTTCCG CGCTTTTCTG 660
 GGTGAACCTA AGACCCAGAT GACATCAGCA GTAAGAGAGC CCAACTACCT TGTTCGGCA 720
 GGATGCTGTA AGGGGTGTGC CTCACCTCTG TGCAACTTCA CTAACTCCAT GGAAGAGAT 780
 CCGCAGACTT CAAGGGAGAT TTTTAATTTT GTACTAAAG CAATTGCTCC TCAGATTGAT 840
 CTGAAGAGAT ATGCTGTGCC CTCAGCTGGC TTGCGCCTAT TTGCCCTGCA TGCACTCTCAG 900
 TTTAGCACTT GCCTTCTGGA CAACTACGTG TCTCTATTG AAGTCTTGT AAAGTGGTGT 960
 GCGCCACAAA ATGAGAAATT GAAAAAAGCT GCACTTTCAG CCCTGGAAATC CTTTCTGAAA 1020
 CAGGTTTCTA ATATGTTGGC GAAAAATGCA GAATGCATA AAAATAAACT CGAGTACTTT 1080
 ATGGAGCAGT TTTATGGAAT CATCAGAAAT GTGGATTGCA ACAACAAGGA GTTATCTATT 1140
 GCTATCCGTC GATATGGACT TTTGTCAGGA CCGTGAAGG TTATAAACGC AAAAGATGTT 1200
 GACTTCATGT ACGTTGAGCT CATTACGCGC TGCAAGCAGA TGTTCCTCAC CCAGACAGAC 1260
 45 ACTGCTGACG ACCGTGTTTA TCAGATGCCA AGCTTCTCC AGTCTGTTGC AAGCGTCTTG 1320
 CTGTACTCTG ACACAGTTC TGAGGTGTAT ACTCCAGTTC TGGAGCACCT CGTGGTGATG 1380
 CAGATAGACA GTTTCCTCCA GTACAGTCCA AAAATGCAGC TGGTGTGTTG CAGAGCCATA 1440
 GTGAAGGTGT TCCTAGCTTT GGCAGCAAAA GGGCCAGTTC TCAGGAATTG CATTAGTACT 1500
 GTGGTGATCT AGGGTTTAAT CAGAATATGT TCTAAACCAAG TGGTCTCTCC AAAGGGCCCT 1560
 50 GAGTCTGAAT CTGAAGACCA CCGTGCTTCA GGGGAAGTCA GAACCTGGCA ATGGAAGGTG 1620
 CCCACATACA AAGACTAGT GATCTCTTC AGACATCTCC TGAGCTCTGA CCAGATGATG 1680
 GATTCATATT TAGCAGATGA AGCATTTTTC TCTGTGAAT CCTCCAGTGA AAGTCTGAAT 1740
 CATTTACTTT ATTTTGAATT TGTAATATCC GTTTTGAAGA TTGTTGAGAA ATTGGATCTT 1800
 ACACCTTGAA TAGACACTGT TGGGGAACAA GAGAATGGAG ATGAGGCGCC TGGTGTTTGG 1860
 55 ATGATCCCAA CTTACAGTCC AGCGGCTAAC TTGCATCCAG CTAAACCTAA AGATTTTTCG 1920
 GCTTTCACTA ACCTGGTGA ATTTTCAGGA GAGATTCTCC CTGAGAAACA AGCAGAAATT 1980
 TTTGAACCAT GGGTGTACTC ATTTTCATAT GAATTAATTT TGCAATCTAC AAGTGTGCCC 2040
 CTCATCAGTG GTTTCTACAA ATTGCTTTCT ATTACAGTAA GAATGCCCAG GAAATATAAA 2100
 TATTTGAGAG GAGTGTAGTCC AAAGAGTCTG AAACACTCTC CTGAAGACCC AGAAAAGTAT 2160
 60 TCTTGTCTTG CTTTATTGTG GAAATTGGC AAAGAGGTGG CAGTTAAAAA GAAGCAGTAC 2220
 AAAGATGAAC TTTTGGCCTC TTGTTTGACC TTTCTCTGT CCTTGGCACA CAACATCATT 2280
 GAATCTGATG TTAGAGCCTA CGTTCCTGCA CTGCAGATGG CTTTCAAACT GGGCCTGAGC 2340
 TATACCCCTT TGGCAGAAGT AGGCCTGAAT GCTCTAGAAG AATGGTCAAT TTATATTGAC 2400
 AGACATGTAA TGCGAGCCTA TTACAAGAGC ATTCTCCCTC GCCTGGATGG ATACCTGAAG 2460
 65 ACTTCAGCCT TGTCAGATGA GACCAAGAAT AACTGGGAAG TGTCACTCTT TTCTCGGGCT 2520
 GCGCAGAAAG GATTTAATAA AGTGGTGTGA AAGCATCTGA AGAAGCAAAA GAACCTTTCA 2580
 TCAACGAAAG CAATATCCTT AGAAGAAATA AGAATTAGAG TAGTACAAAT GCTTGGATCT 2640
 CTAGGAGGAC AAATAAACAA AAATCTTCTG ACAGTCAOGT CCTCAGATGA GATGATGAAG 2700
 AGCTATGTGG CCGTGGAGAG AGAGAAAGCG CTGAGCTTTG CAGTGGCCCT TAGAGAGATG 2760
 70 AAACCTGTCA TTTTCTGGA TGTGTTCCCT CCTCGAGTCA CAGAATTAGC GCTCAGAGCC 2820
 AGTGACAGAC AAATAAAGT TGCAGCCTGT GAATTTTAC ATAGCATGGT TATGTTTATG 2880
 TTGGGCAAG CCACGAGAT GCCAGAAGGG GGACAGGGAG CCCACCCAT GTACCACTC 2940
 TATAAGCGGA CGTTTCTGT GCTGCTTGA CTTGCGTGTG ATGTTGATCA GGTGACAAGG 3000
 CAACTGTATG AGCCACTAGT TATGCACTGT ATTCACTGGT TCACTAACAA CAAGAAATTT 3060
 75 GAAAGTCAGG ATACTGTGTC CTTACTAGAA GCTATATTGG ATGGAATTGT GGACCTGTT 3120
 GACAGTACTT TAAGAGATT TTGTGTCGG TGTATTCSAG AATTCCTTAA ATGTTCCATT 3180
 AAGCAATATA CACCAAGACA GCAGGAGAG AGTCCAGTAA ACACCAATC GCTTTTCAAG 3240
 CGACTTTATA GCCTTGGCTT TCACCCCAAT GCTTTCAAGA GCGTGGGAGC ATCACTTGCC 3300
 TTTAATAATA TCTACAGGGA ATTCAGGGAA GAAGAGTCTC TGGTGAACA GTTGTGTTT 3360
 80 GAAGCCTTGG TGATATACAT GGAGAGTCTG GCCTTAGCAC ATGCAGATGA GAAGTCCCTA 3420
 GGTACAATTC AACAGTGTG TGATGCCATT GATCACCTAT GCCCATCAT TGAAGAAG 3480
 CATGTTTCTT TAAATAAAGC AAAGAAACGA CGTTTGCCGC GAGGATTTCC ACCTTCCGCA 3540
 TCAITGTGTT TATTGGATCT GTTCAAGTGG CTTTLAGCTC ATTGTGGAG GCCCAGACA 3600
 85 GAATGTCGAC ACAAATCCAT TGAACCTTTT TATAAATTCG TTCCTTTATT GCCAGGCAAC 3660
 AGATCCCTTA ATTTGTGGCT GAAAGATGTT CTCAAGGAAG AAGGTGTCTC TTTTCTCATC 3720
 AACACCTTTG AGGGGGGTGG CTGTGGCCAG CCCTCGGGCA TCCTGGCCCA GCCCACCTC 3780
 TTGTACCTTC GGGGGCCATT CAGCCTGCAG GCCACGCTAT GCTGGCTGGA CCGTCTCCTG 3840

	GCOCGTTGG	AGTGCTACAA	CACGTTCAIT	GGCGAGAGAA	CTGTAGGAGC	GCTCCAGGTC	3900
	CTAGGTACTG	AAGCCAGTC	TTCACITTTG	AAAGCAGTGG	CTTTCTTCTT	AGAAAGCATT	3960
	GCCATGCAAT	ACATTATAGC	AGCAGAAAAG	TGCTTTGGCA	CTGGGGCAGC	AGGTAAACAGA	4020
5	ACAAAGCCAC	AAGAGGGAGA	AAGGTACAAC	TACAGCAAAT	GCACCGTTGT	GGTCCGGATT	4080
	ATGGAGTTTA	CCACGACTCT	GCTAAACACC	TCCCGGAGAG	GATGGAAGCT	CCTGAAGAAG	4140
	GACTTGTTGA	ATACACACCT	GATGAGAGTC	CTGGTGACGA	CGCTGTGTGA	GCCCGCAAGC	4200
	ATAGGTTTCA	ACATCGGAGA	CGTCCAGGTT	ATGGCTCATC	TTCCTGATGT	TGTGTGAAT	4260
	CTGATGAAG	CTCTAAAGAT	GTCCCATATC	AAAGATATCC	TAGAGACCCA	TCTGAGAGAG	4320
10	AAATAACAG	CACAGAGCAT	TGAGGAGCTT	TGTGCCGTCA	ACTTGTATGG	CCCTGACGGC	4380
	CAAGTGGACA	GGAGCAGGCT	GGCTGCTGTT	GTGTCTGCCT	GTAAACAGCT	TCACAGAGCT	4440
	GGGCTTCTGC	ATAATATATT	ACCGTCTCAG	TCCACAGATT	TGCATCATTC	TGTTGGCACA	4500
	GAACITCTTT	CCCTGGTTTA	TAAAGGCATT	GCCCTGGAG	ATGAGAGACA	GTGTCTGCCT	4560
	TCTCTAGACC	TCAGTTGTAA	GCAGCTGGCC	AGCGGACTTC	TGGAGTTAGC	CTTTGCTTTT	4620
15	GSAGGACTGT	GTGAGCGCCT	TGTGAGTCTT	CTCCTGAACC	CAGCGTGTCT	GTCCAAGGCG	4680
	TCCTTGGGCA	GCTCACAGGG	CAGGTCATC	CACITCTCCC	ATGGGGAGTA	TTTCTATAGC	4740
	TTGTCTCTAG	AAACGATCAA	CAOGAAATTA	TGAAAAAATC	TGGATCTTGC	TGTATTGGAG	4800
	CTCATGACGT	CTTCAGTGA	TAATACCAAA	ATGGTGAAGT	CCGTTTGTAA	CGGCATGTTA	4860
	GACCAAGAGT	TCAGGAGAGC	AGCAAACCA	AAACACCAAG	GACTGAAACT	TGCGACTACA	4920
20	ATTCTGCAAC	ACTGGAAGAA	GTGTGATTCA	TGGTGGGCCA	AAGATTCCCC	TCTCGAAACT	4980
	AAATATGGCA	TGCTTGCCCT	ACTGGCAAAA	ATTTTACAGA	TGATTCATC	TGTATCTTTT	5040
	AATACAAAGT	ATGGTTCATT	CCCTGAAGTC	TTTACAACAT	ATATTAGTCT	ACTTGCTGAC	5100
	ACAAAGCTGG	ATCTGATTAT	AAAGGGCCAA	GCTGTCACTC	TTCCTTCATT	CTTCACCAGC	5160
	CTCATGAGAG	GCACTCTGGA	GGAACTTAGA	CGTGTCTGGS	AGCAGCTCAT	CGTGTCTCAC	5220
25	TTCCCCATGC	AGTCCAGGGA	ATTTCTCTCA	GGAACTCCGC	GGTTCAATAA	TTATGTGGAC	5280
	TGCATGAAAA	AGTTTCTAGA	TGCATTGGAA	GCCTATGTTT	GTGGAAATTT	GTGGAAATTT	5340
	ATGACAGAA	TTCTTTGTGC	GGAAACAGCAG	CATGTCTATG	AAGAAATATT	TCAATCCAGT	5400
	TTCAGGAGAG	TTGCCAAGAG	GGGTTCAATG	GTCAACAAG	TAGGCTTTCT	GGAAAGCGTG	5460
	TATGAAATGT	TCAGGAAGGA	TGACCCCGCC	CTAAGTTTCA	CACGCCAGTC	CTTTGTGGAC	5520
30	CGCTCCCTCC	TCACCTGCTC	GTGGCACTGT	AGCCTGGATG	CTTTGAGAGA	ATTCTTCAGC	5580
	ACAAATGTGG	TGATGCTGCT	TGATGCTGTT	AAGTCCAGGT	TTACAAAGCT	AAATGAATCT	5640
	ACCTTTGATA	CTCAAAATCA	CAAGAAGATG	GGCTACTATA	AGATTCTAGA	CGTGATGTAT	5700
	TCTGCGCTGT	CCAAAGATGA	TGTTTATGCT	AAGGAATCAA	AAATTAATCA	AGTTTTCAT	5760
	GGCTCGTGT	TTACAGAAGG	AAATGAACCT	ACAAAGACAT	TGATTAAATT	GTGCTACGAT	5820
35	GCATTTACAG	AGAACATGCG	AGGAGAGAAT	CAGCTGCTGG	AGAGGAGAAG	ACTTTACCAT	5880
	TGTGACAGCT	ACAACTGCGC	CATATCTGTC	ATCTGCTGTG	TCTTCAATGA	GTTAAATTTT	5940
	TACCAAGGTT	TTCTGTTTAT	TGAAAAACCA	GAAGAAGACT	TGCTTATTTT	TGAAAACTG	6000
	ATCGACCTGA	AGCGCCGCTA	TAATTTTCTT	GTAGAAGTTG	AGGTTCTCTAT	GGAAAGAAAG	6060
	AAAAAGTACA	TTGAAATTTAG	GAAGAAGGCC	AGAGAAGCAG	CRAATGGGGA	TTGAGATGCT	6120
40	CCCTTCTATA	TGCTTCTCCCT	GTCAATTTTG	GCAGACAGTA	CCCTGAGTGA	GGAAATGAGT	6180
	CAATTTGATT	TCTCAACCGG	AGTTTCAAGC	TATTCATACA	GCTCCCAAGA	CCCTAGACCT	6240
	GCCACTGGTC	GTCTTCCGAG	ACGGGAGCAG	CGGGACCCCA	CGGTGCATGA	TGATGTGCTG	6300
	GAGCTGGAGA	TGGAAGAGCT	CAATCGGCAT	GAGTGCATGG	CGCCCTGACG	GGCCCTGGTC	6360
	AAGCACATGC	ACAGAAGCCT	GGGCCCCGCT	CAAGGAGAAG	AGGATTCACT	GCCAAAGAT	6420
45	CTTCTCTTCT	GGATGAAAT	CCTCCATGGC	AACTGCGGAA	ATCCAAATAG	ACCAATTAAT	6480
	ATCGTCTCT	TCTTAGCCAA	GCTTGTATT	AATACAGAAG	AGGTCTTCTG	CCCTTACGGG	6540
	AAGCACTGGC	TTAGCCCTTT	GCTGCAGCTG	GCTGCTCTTG	AAAACAATGG	AGGAGAAGGA	6600
	ATTTACTACA	TGGTGGTTGA	GATAGTGCC	ACTATTCTTT	CATGGACAGG	CTTGGCCACT	6660
	CCACAGGGGG	TCCCTAAAGA	TGAAGTGTTA	GCAAAATGAT	TGCTTAATTT	CCTAATGAAA	6720
50	CATGTCTTTC	ATCCAAAAG	AGCTGTGTTT	AGACACAACC	TTGAAATTTAT	AAAGACCCCT	6780
	GTGAGTGTCT	GGAGAAAGT	TTTATCCATC	CCTTATAGGT	TAATATTGTA	AAAGTTTTC	6840
	GGTAAAGATC	CTAATTTCAA	AGACAACCTA	GTAGGGATTC	AATGTCTAGG	CATGTGATG	6900
	GCCAAATGACC	TGCTTCCCTA	TGACCCACAG	TGTGGCATCC	AGAGTAGCGA	ATACTTCCAG	6960
55	GCTTTGGTGA	ATAATATGTC	CTTTGTAAAG	TATAAAGAAG	TGTATGCCGC	TGCAGCAGAA	7020
	GTCTTAGGAG	TTATACTTGG	ATATGTTATG	GAGAGAAAAA	ACATACTGGA	GGAGTCTCTG	7080
	TGTGAACCTG	TTGCGAAACA	ATTGAAGCAA	CATCAGAATA	CTATGGAGGA	CAAGTTTATT	7140
	GTGTGCTTGA	ACAAAGTGAC	CAAGAGCTTC	CCTCTCTTGG	CAGACAGGTT	CATGAATGCT	7200
	GTGTTCTTTC	TGCTGCCAAA	ATTTTATGGA	GTGTTGAAAA	CACCTCTGCT	GGAGGTGGTA	7260
60	CTTTGTCGTG	TGGAGGGAAT	GACAGAGCTG	TACTTCCAGT	TAAAGAGCAA	GGACTTCTGT	7320
	CAAGTCATGA	GACATAGAGA	TGATGAAAGA	CAAAAAGTAT	GTTTGGACAT	AATTTATAAG	7380
	ATGATGCCAA	AGTTTAAACC	AGTAGAATCT	CGAGAATCTC	TGAACCCCGT	TGTGGAATTC	7440
	GTTCGCCATC	CTTCTACAAC	ATGTAGGGAA	CAAAATGTATA	ATATTCTCAT	GTGGATTCAAT	7500
	GATAATTACA	GAGATCCAGA	AAGTGAGACA	GATAATGACT	CCCAGGAAAT	ATTTAAGTTG	7560
65	GCAAAAGATG	TGCTGATTCA	AGGATTGATC	GATGAGAAC	CTGGACTTCA	ATTAATTAAT	7620
	CGAAATTTCT	GGAGCCATGA	AACATAGGTT	CCTTCAAATA	CCTTGGACCG	GTGTCTGGCA	7680
	CTAAATTCCT	TATATTCTCC	TAAGATAGAA	GTGCACITTT	TAAGTTTAGC	AACAAATTTT	7740
	CTGCTCGAAA	TGACAGCAT	GAGCCAGAT	TATCCAAACC	CCATGTTTGA	GCATCTCTG	7800
	TCGAATTCGG	AAITTCAGGA	ATATACCAAT	GATTCGTATT	GGCGTTTCCG	AAGTACTGTT	7860
70	CTCACTCCGA	TGTTTGTGGA	GACCCAGGCC	TCCAGGGGCA	CTCTCCAGAC	COGTACCCAG	7920
	GAAGGGTCCC	TCTCAGCTCG	CTGGCCAGTG	GCAGGGCAGA	TAAGGGCCAC	CCAGCAGCAG	7980
	CATGACTTCA	CACGTGACAC	GACTGCAGAT	GGAGAAGGCT	CATTGTATTG	GCTGACCGGG	8040
	AGCAGCACTG	ACCCGCTGCT	CGACCAACCC	AGTCCCTCAT	CTGACTCCTT	GCTGTTTGCC	8100
	CACAAGAGGA	GTGAAAGGTT	ACAGAGAGCA	CCCTTGAAGT	CAGTGGGGCC	TGATTTTGGG	8160
75	AAAAAAGAGC	TGGCCCTTCC	AGGGGACGAG	GTGGATAACA	AAGTGAAAGG	TGCGCCGGG	8220
	CGGACGGACC	TACTACGACT	CGCGAGACGG	TTTATGAGGG	ACCAGGAGAA	GCTCAGTTTG	8280
	ATGTATGCCA	GAAAAGGCGT	TGCTGAGCAA	AAACGAGAGA	AGGAAATCAA	GAGTGAAGTTA	8340
	AAAAATGAAG	AGGATGCCCA	GGTCTTCTG	TACAGAAGCT	ACCGGCAAGG	AGACCTTCT	8400
	GACATTTCAG	TCAAGCACAG	CAGCCTCATC	ACCCGTTTAC	AGGCGTGGC	CCAGAGGGAC	8460
	CCAAATATTG	CAAAAACAGT	CTTTAGCAGC	TTGTTTCTG	GAATTTTGA	AGAGATGGAT	8520
80	AAATTTAAGA	CACGTGCTGA	AAAAACAAC	ATCACTCAA	AGTTGCTTCA	AGACTTCAAT	8580
	CGTTTCTTFA	ATACCACTTT	CTCTTCTTT	CCACCTTTG	TCTCTGTAT	TGAGGACATT	8640
	AGCTGTGACG	ACGCAGCCCT	GCTGAGCCTC	GACCCAGCGG	CTGTTAGCGC	TGTTTGCTG	8700
	GCCAGCCTAC	AGAGCCCGGT	GGGCATCCGC	CTGCTAGAGG	AGGCTCTGCT	CGCCCTGCTG	8760
	CCTGCTGAGC	TGCTGCCCAA	GCGAGTCCGT	GGGAAGGCCC	GCTCCCTTCC	TGATGCTCTC	8820
	AGATGGGTGG	AGCTTGCTAA	GCTGTATAGA	TCAATTGGAG	AATACGACGT	CCTCCGTGGG	8880
85	ATTTTTACCA	GTGAGATAGG	AACAAAGCAA	ATCACTCAGA	GTGCATTATT	AGCAAGAGCC	8940
	AGAAGTGATT	ATTCTGAAGC	TGCTAAGCAG	TATGATGAGG	CTCTCAATAA	ACAAGACTGG	9000
	GTAGATGGTG	AGCCCCACAG	AGCCGAGAAG	GATTTTGGG	AACITGACATC	CCTTGACTGT	9060

	TACAACCCACC	TTGCTGAGTG	GAAATCACTT	GAATACTGTT	CTACAGCCAG	TATAGACAGT	9120
	GAGAACCCCC	CAGACCTAAA	TAAAATCTGG	AGTGAACCAT	TTTATCAGGA	AACATATCTA	9180
	CCTTACATGA	TCCGCAGCAA	GCTGAAGCTG	CTGCTCCAGG	GAGAGGCTGA	CCAGTCCCTG	9240
5	CTGACATTTA	TTGACAAGG	TATGCAAGG	GAGCTCCAGA	AGGCGATTCT	AGAGCTTCAT	9300
	TACAGTCAAG	AGCTGAGTCT	GCTTACCTC	CTGCAAGATG	ATGTTGACAG	AGCCAAATAT	9360
	TACATTCAAA	ATGGCATTCA	GAGTTTATG	CAGAAITATT	CTAGTATTGA	TGTCCTCTTA	9420
	CACCAAGATG	GACTCACCAA	ATTGCAGTCT	GTACAGGCTT	TAACAGAAAT	TCAGGAGTTC	9480
	ATCAGCTTTA	TAAGCAAAAC	AGGCAATTTA	TCATCTCAAG	TTCCCTTTAA	GAGACTTCTG	9540
10	AACACCTGGA	CAACAGATA	TCCAGATGCT	AAAATGGACC	CAATGAACAT	CTGGGATGAC	9600
	ATCATCACAA	ATCGATGTTT	CTTTCTCAGC	AAAATAGAGG	AGAAGCTTAC	CCCTCTTCCA	9660
	GAGATATAA	GTATGAATGT	GGATCAAGAT	GGAGACCCCA	GTGACAGGAT	GGAGATGCAA	9720
	GAGCAGGAAG	AAGATATCAG	CTCCCTGATC	AGGAGTTGCA	AGTTTTCCAT	GAAAATGAAG	9780
	ATGATAGAGA	GTGCCCGGAA	GCAGAACAA	TTCTCACTTG	CTATGAAACT	ACTGAAGGAG	9840
15	CTGCATAAAG	AGTCAAAAAC	CAGAGACGAT	TGGCTGGTGA	GCTGGGTGCA	GAGCTACTGC	9900
	CGCCTGAGCC	ACTGCCGGAG	CGGTGCCAG	GGCTGCTCTG	AGCAGGTGCT	CACCTGTGCTG	9960
	AAAAACAGTCT	CTTTGTTGGA	TGAGAACAA	GTGTCAAGCT	ACTTAAGCAA	AAATATTCTG	10020
	GCTTTCCGTT	ACCAAGCAAT	TCTCTGGGT	ACAACTTACA	GGATCATAGC	GAATGCTCTC	10080
	AGCAGTGAGC	CAGCCTGCCT	TGCTGAAATC	GAGGAGGACA	AGGCTAGAAG	AATCTTAGAG	10140
20	CTTTCTGGAT	CCAGTTTACA	GGATTTCAGG	AAGGTGATCG	CGGCTCTGTA	CCAGAGAGCA	10200
	TTCCAGCACC	TCTCTGAGGC	TGTGCAGGCG	GCTGAGGAGG	AGGCCAGGCC	TCCCTCTCTG	10260
	AGCTGTGGGC	CTGCAGCTGG	GGTGATTGAT	GCTTACATGA	CGCTGGCAGA	TTTCTGTGAC	10320
	CAACAGCTGC	GCAAGGAGCA	AGAGAATGCA	TCAGTTATTG	ATTCTGCAGA	ACTGCAGGCG	10380
	TATCCAGCAC	TTGTGTGTGA	GAAAATGTTG	AAAGCTTTAA	AATTAATTC	CAATGAAGCC	10440
25	AGATTGAAGT	TTCTTAGATT	ACTTCAGATT	ATAGAACCGT	ATCCAGAGGA	GACTTTGAGC	10500
	CTCATGACAA	AAGAGATCTC	TTCCGTTCCC	TGCTGGCAGT	TCATCAGCTG	GATCAGCCAC	10560
	ATGGTGGCCT	TACTTGACAA	AGACCAAGCC	GTGCTGTTCT	AGCACTCTGT	GGAGAAATTC	10620
	ACTGATAACT	ACCCGACAGC	TATTGTTTAT	CCCTTCATCA	TAAGCAGCGA	AAGCTATTCC	10680
	TTCAAGGATA	CTTCTACTGG	TCATAAGAAT	AAGGAGTTTG	TGGCAAGGAT	TAAAAGTAA	10740
30	TTGGATCAAG	GAGGAGTGAT	TCAAGATTTT	ATTAATGCCCT	TAGATCAGCT	CTCTAATCCT	10800
	GAACTGCTCT	TAAAGATGAT	GAGCAATGAT	GTAAAGCTG	AAGTACAGAA	AACCCCTGTA	10860
	AATAAAAAAA	ACATTGAAA	AATGTATGAA	AGAAATGTATG	CAGCCTTGGG	TGACCCCAAG	10920
	GCTCCAGGCC	TGGGGCCTTT	TAGAAGGAAG	TTTATTTCAGA	CTTTTGGAAA	AGAATTTGAT	10980
	AAACATTTTG	GGAAAGGAGG	TTCTAAACTA	CTGAGAAATGA	AGCTCAGTGA	CTTCAACGAC	11040
35	ATTACCAACA	TGCTACTTTT	AAAAATGAAC	AAAGACTCAA	AGCCCCCTGG	GAATCTGAAA	11100
	GAACTGTTAC	CTTGAGTATG	CGACTTCAA	GTGGAGTTCC	TGAGAAATGA	GCTGGAGATT	11160
	CCCGTCTAGT	ATGACGGTAG	GGGAAAGCCA	TTGCCAGAGT	ACCAAGCTGG	AATCGCCGGG	11220
	TTTATGAGC	GGGTGACAGT	CATGGCTCT	CTGGGAAGGC	CCAAGCGCAT	CATCATCCGT	11280
	GGCCATGAGC	AGAGGGAACA	CCCTTTCTCT	GTGAAGGGTG	GCGAGGACCT	GCGGCAGGAC	11340
40	CAGCGCGTGG	AGGACGCTCT	CCAGGTCATG	AATGGGATCC	TGGCCCAAGA	CTCGCCCTGC	11400
	AGCCAGAGGG	CCCTGAGCTG	GAGGACCTAT	AGCGTTGTGC	CCATGACCTC	CAGGTTAGGA	11460
	TTAATTTAGT	GGCTTGAAA	TACTGTTACC	TTGAAGGACC	TTCTTTTGAA	CACCATGTCC	11520
	CAAGAGGAGA	GGGCGCTTAA	CCTGAGTGAT	CCCAGGGCAC	CGCCGTGTGA	ATATAAAGAT	11580
	TGGCTGACAA	AAATGTCTAG	AAAACATGAT	GTTGGAGCTT	ACATGCTAAT	GTATAAGGGC	11640
45	GCTAATCTGA	CTGAAACAGT	CAOGTCTTTT	AGAAAACGAG	AAAGTAAAGT	GCCTGTCTGAT	11700
	CTCTTAAAGC	GCGCCTTCCT	GAGGATGAGT	ACAAGCCCTG	AGGCTTTCTC	GCGCCTCCCG	11760
	TCCCATCTGG	CCAGCTCTCA	CGCTCTGATA	TGCATCAGCC	ACTGGATCCT	CGGGATTGGA	11820
	GACAGACATC	TGAACAACTT	TATGGTGGCC	ATGGAGACTG	GCGGCGTGAT	CGGGATCGAC	11880
	TTTGGGCAATG	CGTTTGGATC	CGCTACACAG	TTTCTGCCAG	TCCCTGAGTT	GATGCCCTTT	11940
50	CGGCTAACTC	GCCAGTTTAT	CAATCTGATG	TTACCAATGA	AAGAAACGGG	CCTTATGTAC	12000
	AGCATCATGG	TACACGCACT	CCGGGCTTTC	CGCTCAGACC	CTGGCCTGCT	CACCAACACC	12060
	ATGGATGTGT	TTGTCAAGGA	GCCCTCCTTT	GATTGGAAAA	ATTTTGAACA	GAAAATGCTG	12120
	AAAAAGGAGG	GGTCACTGAT	TCAAGAAATA	AATGTTGCTG	AAAAAAATG	GTACCCCGGA	12180
	CAGAAAATAT	GTTACGCTAA	GAGAAAGTTA	GCAGGTGCCA	ATCCAGCAGT	CATTACTTGT	12240
	GATGAGCTAC	TCTCTGGTCA	TGAGAAAGCC	CCTGCCTTCA	GAGACTATGT	GGCTGTGGCA	12300
55	CGAGGAAGCA	AAGATCAACA	CATTGCTGCC	CAAGAACCCG	AGAGTGGGCT	TTTCAAGAGG	12360
	ACTCAAGTGA	AGTGCCTGAT	GGACCAAGCA	ACAGACCCCA	ACATCCTTGG	CAGAACTCTG	12420
	GAAGGATGGG	AGCCCTGGAT	GTGAGGCTCT	TGGGAGTCTG	CAGATAGAAA	GCATTACATT	12480
	GTTTAAAGAA	TCTACTATAC	TTTGGTTGGC	AGCATTCCAT	GAGCTGATT	TCCTGAAACA	12540
60	CTAAAGAGAA	ATGCTTTTGG	TGCTACAGTT	TCGTAGCATG	AGTTTAAATC	AAGATTATGA	12600
	TGAGTAAATG	TGATGGGTTT	AAATCAAAGA	TAAGGTTATA	GTAACATCAA	AGATTAGGTG	12660
	AGGTTTATAG	AAAGATAGAT	ATCCAGGCTT	ACCAAAGTAT	TAAGTCAAGA	ATATAATATG	12720
	TGATCAGCTT	TCAAAGCAAT	TACAAGTGCT	GCAAGTTAGT	GAAACAGCTG	TCTCCGTAAA	12780
	TGGAGGAATG	GTGGGGAAGC	CTTGGAAATG	CCTTCTGGTT	CTGGCACATT	GGAAAGCACA	12840
65	CTCAGAGGCG	TTTCATCACA	AGATTTTGGG	AGAGTAAAGC	TAAGTATAGT	TGATGTAAAC	12900
	TTGTAGAAGC	AGCATAGGAA	CAATAAGAAC	AATAGGTAAA	GCTATAATTA	TGGCTTATAT	12960
	TTAGAAATGA	CTGCATTGTA	TATTTTAGGA	TATTTTCTA	GGTTTTTCTC	TTTCAATTTA	13020
	TTCTCTTCTA	GTTTTGACAT	TTTATGATAG	ATTTGCTCTC	TAGAAGGAAA	CGCTTTTATT	13080
	TAGGAGGGCA	AAAATTTTGG	TCATAGCATT	CACTTTGTCT	ATTCCAAATC	ACAACTGGAA	13140
70	GATACATAAA	AGTGCTTTGC	ATTGAATTTG	GGATAACTTC	AAAAATCCCA	TGGTTGTTGT	13200
	TAGGGATAGT	ACTAAGCATT	TCAGTTCCAG	GAGAATAAAA	GAAATTCCTA	TTTGAATGTA	13260
	ATTCTCTCAT	TGGAGGAAAA	AAAGCATGCA	TTCTAGCACA	ACAAGATGAA	ATTATGGAAT	13320
	ACAAAAGTGG	CTCCTTCCCA	TGTGCAGTCC	CTGTCCCCCC	CCGCCAGTCC	TCCACACCCA	13380
	AACGTGTTCT	GATGTGCTTT	TAGCTTTTGG	TTGTTTTTTT	TTTTCTTCTC	AACACTTGTA	13440
75	TTTGAGGCTC	CTTCTGTGAT	TTTGAGAAGT	ATACTCTTGA	GTGTTTAAAT	AAGTTTTTTT	13500
	CCAAAAGTA						

Seq ID NO: 99 Protein sequence:
Protein Accession #: NP_008835.5

80	1	11	21	31	41	51	
	MAGSGAGVRC	SLRLQETLS	AADRCGAALA	GHQLIRGLGQ	ECVLSSSPAV	LALQTSLVFS	60
	RDPLLVFVR	KSLNLSFPR	CREELKFLC	IFLEKMGQRI	APYSVEIKNT	CTSVYTKDRA	120
85	AKCKIPALDL	LILKLTQFRS	SRLMDEFKIG	ELPSKPYGEL	ALKKKIPDTV	LEKVYELLGL	180
	LGEVHPSEMI	NNAENLFRAP	LGELKTQMTS	AVREPFLFVL	AGCLKGLSSL	LCNPTKSMEE	240
	DPQTSREIFN	FVLKAIRPQI	DLKRYAVPSA	GLRLPALHAS	QFSTCLLDNY	VSLFEVLLKW	300
	CAHTNVELKK	AALSALSF	KQVSNMVARN	AEHKNKLQY	FMEQFYGIIR	NVDSNNKELS	360

	IAIRGYGLFA	GPCKVINAXD	VDFMYVELIQ	RCKQMFLTQT	DTGDDRVYQM	PSFLQSVASV	420
	LLVLDTPVEV	YTPVLEHLVV	MQIDSFPQYS	PRMQLVCCRA	IVKVFLALAA	KGFVLRNCIS	480
	TVVHQGLIRI	CSKPVVLPKG	PESESEDHRA	SGEVRTGKWK	VPTYKDYVDL	FRHLLSSDQM	540
	MDSILADEAF	PSVNSSESLS	NHLLYDEFVK	SVLKIVEKLD	LTLLEIQTVE	QENGDEAPGV	600
5	WMIPSTDPA	NLHPAKPKDF	SAFINLVEPC	REILPEKQAE	FFEPWVYSFS	YELILQSTRL	660
	PLISGFYKLL	SITVRNAKKI	KYFEGVSPKS	LKHSPEDEPK	YSCFALFVKF	GKEVAVRMKQ	720
	YKDELLASCL	TFLLSLPHNI	IELDVRAYVP	ALQMAPKLGL	SYTPLAEVGL	NALEWSIYI	780
	DRHVMQPYK	DILPCLDGYL	KTSALSDETK	NNWEVSALSR	AAQKGFNKVV	LKHLKCTKNL	840
	SSNEAISLEE	IRTRVQMLG	SLGGQINKNL	LTVTSSDEMM	KSYVAMDREK	RLSPAVFPRE	900
10	MKPVIFLDVF	LPRVTEALAT	ASDRQTKVAA	CELLHSMVMF	MLGKATQMPF	GGQGAPPMYQ	960
	LYKRTFPVLL	RLACDQDVOT	RQLYEPLVMQ	LIHWFTNNKK	FESQDTVALL	EAILEDGIVDP	1020
	VDSTLRDFCG	RCIREPLKWS	IKQITPQQOE	KSPVNTKSLF	KRLYSLALHP	NAFKRLGASL	1080
	AFNNIYREFR	EESLVEQFV	FEALVIYMES	LALAHADKES	LGTIQCCDA	IDHLCRIIEK	1140
	KHVSANAKK	RRLPRGFPFS	ASLCLLDLVK	WLLAHCGRPQ	TECRHKSIEL	FYKFPVLLPG	1200
15	NRSPLNLKLD	VLEKEGVSLF	INTFEGGCGG	QPSGILAQPT	LLYLRGPPSL	QATLCWLDLL	1260
	LAALCYNTFF	IGERTVGAIG	VLGTEAQSSL	LKAVAFPLES	IAMHDIIAAE	KCFGTGAAGN	1320
	RTSPQEGERY	NSYKCTVVVR	IMEFTTLLN	TSPEGWKLK	KDLNTHLMR	VLVQTLCEPA	1380
	SIGFNIQDQV	VMAHLPDVCV	NLMKALKMSP	YKDIETHLR	EKITAQSTEE	LCAVNLYGPD	1440
	AQVDRSRLAA	VVSACKQLHR	AGLLHNILPS	QSTDLLHVSQ	TELLSLVYKG	IAPGDERQCL	1500
20	PSLDLSCQPL	ASGLLELAPA	PGLCERLVS	LLLNPAVLST	ASLGSSQGSV	IHFSGHEYFY	1560
	SLFSETINTE	LLKNLDAVL	ELMQSSVDNT	KMVSAVLNGM	LDQSFREERAN	QKHQGLKAT	1620
	TILQHWKCD	SWAKEDSPLE	TKMAVLALLA	KILOIDSSVS	FNTSHGSPFE	VFTTYISLLA	1680
	DTKLDLHLRG	QAVTLFPFFT	SLTGGSLLEL	RRVLEQLIVA	HFPMQSREFP	PCTPRPNYV	1740
	DCMKFLDAL	ELSQSPMLLE	LMTFVLCREQ	QHVMEELPQS	SFRRIARRGS	CVTQVGLLES	1800
25	VYEMFRKQDP	RLSFTKQSFV	DRSLTLTLNH	CSLDALREFP	STIVVDAIDV	LKSRFTKNE	1860
	STPDQTITKG	MGYKILQDM	YSRLPKDDVH	AKESKINQVF	HGSCITEGNE	LTKTLIKLCY	1920
	DAPTENMAGE	NQLEERRRIL	HCAAYNCAIS	VICCVFNELK	FYQGFLESEK	PEKNLLIFEN	1980
	LIDLKRRYNF	PVEVEVPMER	KKKYIEIRKE	AREAANGSDS	GPSYMSLSY	LADSTLSEEM	2040
	SQDFSTGVQ	SYSSYSQDPR	PATGRFRRE	QRDPTVHDDV	LELEMDLNR	HECMAPLTAL	2100
30	VKMHRSLQPL	PQGEEDSVPR	DLPSMMKFLH	GKLGNPVPL	NIRLFLAKLV	INTTEVFRPY	2160
	AKHWSPLLL	LAASENNGGE	GIHYMVVEIV	ATILSWTGLA	TPTGVPKDEV	LANRLNLFML	2220
	KHVFHFKRAV	FRNLEILITL	LVECKWDCLS	IPYRLIFEKP	SGKDPSNKN	SVGILQLGIV	2280
	MANDLPPYDP	QCGIQSSEYF	QALVNMMSFV	RYKEVYAAAA	EVLGLILRYV	MERNKILEES	2340
35	LCBLVAKQLK	QHONTMEDKF	IVCLNKVTKS	FPLADRFPM	AVFFLLPKFH	GVLLKTLCEV	2400
	VLCRVEGMTF	LYQLQSKDPD	VQVMHRDDE	RQKVCLEIY	KMPKLPKPE	LRELINPVVE	2460
	FVSHPTTCTR	EQMYNLLMWI	HDNYRDPSE	TDNDSQEIFK	LAKDVLIQGL	IDENPGLQLI	2520
	IRNFWSHETR	LPSTNLDRLT	ALNSLYSPRI	EVHFLSLATN	FLEMTSMSP	DYPNFMFEHP	2580
	LSECEPQEYT	IDSDFRFRST	VLTPMFVETQ	ASQGTQLQRT	QEGSLSARWP	VAGQIRATQQ	2640
40	QHDFTLTQTA	DGRSSDFDWT	GSSTDPVLDH	TSPSSDLSLF	AHKRSERLQR	APLKSVPDPF	2700
	GKRLQLPGDP	EVNKNVKGAA	GRDILLRLRR	RFMRDQEKLS	LMYARKGVAE	OKREKEIKSE	2760
	LKMKQDAQV	LYRSYRHGDL	FDIQIKHSSL	ITPLQAVQAR	DPHAKQLFS	SLFSGILKEM	2820
	DKPKTLSEKN	NITQKLQDP	NRLNNTTFSF	FPPFVSCIQD	ISCQHAALLS	LDPAAVSAGC	2880
	LASLQPVGVI	RLLEEALLRL	LPAELPAKRV	RGKARLPDDV	LRWVELAKLY	RSIGEYDVL	2940
45	GIFTSEIGTK	QITQSALLAE	ARSDYSEAAK	QYDEALNKQD	WVDGEPTBAE	KDFWELASLD	3000
	CYNHLAEWKS	LEVCSFTASID	SENPPDLNKI	WSEPFYQETY	LPYMIKSLK	LLQGEADQS	3060
	LLTFIDKAMH	GELQKALIEL	HYSQELSLLY	LLQDDVDRAK	YYIQNGIQSF	MQNYSSIDVL	3120
	LHQSRLLTKL	SQVALTEIQE	FISFISKQGN	LSSQVPLKRL	LNTWNTNRPD	AKMDPMNIWD	3180
50	DIITNRCFFL	SKIEEKLTP	PEDNSMNVQD	DGDPDSRMEV	QEEDIEDSL	IRSCKPSMKM	3240
	KMIDSRKQKN	NFSLAMKLLK	ELHKEKTRD	DWLVSQVQSY	CRLSHCRSRS	QCSSEQVLT	3300
	LKTVSLLEDN	NVSSYLSKNI	LAFRDQNIL	GTTYRIIANA	LSSEPACLA	IEEDKARRIL	3360
	ELSGSSSEDS	EKVIAGLYQR	APQHLSEAVQ	AAEEEAQPPS	WSCGPAAGVI	DAYMTLADFC	3420
	DQQLRKEEN	ASVIDSAELQ	AYPALVVEIM	LKALKLNSNE	ARLKFPRLLQ	IERYPEESTL	3480
55	SLMTKEISSV	PCNQFISWIS	HMVALLDKDQ	AVAVQHSVEE	ITDNYPQAI	YFPIISSESY	3540
	SFEDTSTGHK	NKEFVARIKS	KLDQGGVQID	FINALDQLSN	PELLFKDWSN	DVRAELAKTP	3600
	VNKNIEKMY	ERNYAAALGDP	KAPGLGAFRR	KFIQTGKEF	DKHFGKGGSK	LLRMKLSDFN	3660
	DITNMLLLKM	NKDSKPPGNL	KECSPHMSDF	KVEFLRNELE	IPGQYDGRGK	PLPEYHVRIA	3720
	GFDERVTMA	SLRRPKRIII	RGHDEREHPF	LVKGGEDLRQ	DQRVEQLFQV	MNGILAQDSA	3780
60	CSQRALQTRT	YSVVPMTSRL	GLIEWLENTV	TLKDLLNTM	SQEEKAAYLS	DPRAPPCEYK	3840
	DWLTKMSGKH	DVGAYMLMYK	GANKTETVTS	FRKRESKVA	DLLKRAFVRM	STSPEAFIAL	3900
	RSHFASBAL	ICISHWILGI	GDRHLNFMV	AMETGGVIGI	DFGHAFGSAT	QFLPVPELMP	3960
	FRITRQFINL	MLPMKGTGLM	YSIMVHALRA	FRSDPGLLTN	TMDVPVKEPS	FDWKNFQKMY	4020
	LKKGGSWQIE	INVAEKWYWP	RQKICYAKRK	LAGANPAVIT	CDELLLGHEK	APAFRDYVAV	4080
	ARGSKDHNR	AGSPESGLSE	ETQVKCLMDQ	ATDFNLLGRT	WEGWEPWM		

Seq ID NO: 100 DNA sequence
Nucleic Acid Accession #: NM_000673
Coding sequence: 101-1225

70	1	11	21	31	41	51	
	ATGTGAAGGC	ACAAGCTGCT	GTTATATACA	ACAGAGTGAA	CTGAGCATCA	GTCAGAAAAA	60
	GTCATGATTT	GCAGAAATAC	AGATCCCAAG	CAAGACAGG	ATGGGCACTG	CTGGAAGAA	120
	TATTAAATGC	AAAGCAGCTG	TGCTTTGGGA	GCAGAAGCAA	CCCTTCTCCA	TTGAGGAAAT	180
	AGAAGTGCC	CCACCAAGAA	CTAAAGAAAT	TCGCATTAA	ATTTTGCCCA	CAGGAATCTG	240
75	TCGCACAGAT	GACCATGTGA	TAAAAGGAAC	AATGGTGTC	AAGTTTCCAG	TGATTGTGGG	300
	ACATGAGGCA	ACTGGGATG	TAGAGAGCAT	TGGAGAAGGA	GTGACTACAG	TGAAACCAGG	360
	TGACAAAGTC	ATCCCTCTCT	TTCTGCCACA	ATGTAGAGAA	TGCAATGCTT	GTCGCAACCC	420
	AGATGGCAAC	CTTTGCATT	GGAGGATAT	TACTGGTCGT	GGAGTACTGG	CTGATGGCAC	480
	CACCAAGATT	ACATGCAGG	GCAACCAAGT	ACACCACTTC	ATGAACACCA	GTACATTAC	540
80	CGAGTACACA	TGGTGGATG	AATCTTCTGT	TGCTAAGATT	GATGATGCAG	CTCTCTCTGA	600
	GAAAGTCTGT	TTAATTTGGT	GTGGGTTTTT	CACTGGATAT	GGCGCTGCTG	TTAAACTTGG	660
	CAAGGTCAAA	CCTGGTTTCA	CTTGGTCTGT	CTTGGGCTGT	GGAGGAGTTG	GCCTGTGAGT	720
	CATCATGGGC	TGTAGATCAG	CTGGTGCATC	TAGGATCATT	GGGATTGACC	TCAACAAAGA	780
	CAAAATTGAG	AAGGCCATGG	CTGTAGGTGC	CACTGAGTGT	ATCAGTCCCA	AGGACTCTAC	840
85	CAAAACCATC	AGTGAGGTGC	GTGCAGAAAT	GACAGGCAAC	AACGTGGGAT	ACACCTTTGA	900
	AGTTATTGGG	CATCTTGAAT	CCATGATTGA	TGCCCTGGCA	TCCTGCCACA	TGAACATATG	960
	GACCAAGCTG	GTGTGAGGAG	TTCTCTCATC	AGCCAAGATG	CTACCTATG	ACCGATGATT	1020

5
10
15
20

GCTCTTCACT GGACGCACAT GGAAGGGATG TGTCTTTGGA GGTTTGAAAA GCAGAGATGA 1080
 TGTCCCAAAA CTAGTGACTG AGTTCCCTGGC AAAGAAATTT GACCTGGACC AGTTGATAAC 1140
 TCATGTTTAA CCATTTAAAA AAATCAGTGA AGGATTTGAG CTGCTCAATT CAGGACAAA 1200
 CATTGGAAGC GTCTGACGT TTTGAGATCC AAAGTGGCAG GAGGTCTGTG TTGTCATGGT 1260
 GAACGTGAGT TTCTCTGTG AGAGTTCCCT CATCTGAAAT CATGTATCTG TCTCACAAT 1320
 ACAGACATAA GTAGAAGATT TGTGAAGAC ATAGAACCCT TATAAGAAT TATTAACCTT 1380
 TATAAACATT TAAAGTCTTG TGAGCACCTG GGAATTAGTA TAATAACAAT GTTAATATTT 1440
 TTGATTACA TTTTGTAAAG CTATAATTGT ATCTTTTAAG AAAACATACA CTGGGATTTC 1500
 TATGTTGAAA TGGAGATTTT TAAGAGTTT AACCAGCTGC TGCAGATATA TAACTCAAAA 1560
 CAGATATAGC GTATAAAGAT ATAGTAAATG CATCTCCAG AGTAATATTC ACTTAACACA 1620
 TTGAAACTAT TATTTTITAG ATTTGAATAT AAATGTATTT TTTAAACACT TGTATGAGT 1680
 TAACCTGGAT TACATTTTGA AATCAGTTCA TTCCATGATG CATATTACTG GATTAGATTA 1740
 AGAAAGACAG AAAAGATTAA GGGACGGGCA CATTTTTCAA CGATTAAGAA TCATCATTAC 1800
 ATAACTTGGT GAACATGAAA AAGTATATCA TATGGGTACA CAAGGCTATT TGCCAGCATA 1860
 TATTAATATT TTAGAAAATA TTCCTTTTGT AATACTGAAT ATAAACATAG AGCTAGAGTC 1920
 ATATTATCAT ACTTATCATA ATGTTCAATT TGATACAGTA GAATTGCAAG TCCCTAAGTC 1980
 CCTATTCACT GTGCTTAGTA GTGACTCCAT TTAATAAAAA GTGTTTITAG TTTTAAACAA 2040
 CTAACCCG

Seq ID NO: 101 Protein sequence:
 Protein Accession #: NP_000664

25
30

1 11 21 31 41 51
 | | | | |
 MGTAGKVIK KAAVLWEQKQ PPSIEEIEVA PPKTKEVRIK ILATGICRTD DHVIKGTMTVS 60
 KFPVIVGHEA TGIVESIGEG VTTVKPGDKV IPLFLPQCRE CNACRNPQGN LCIRSDITGR 120
 GVLADGTTTF TCKGKPVHFF MNTSTFTFT VVDESSVAKI DDAAPPEKVC LIGCGFSTGY 180
 GAUVKTKGVK PGSTCVVFLG GGVGLSVIMG CKSAGASRII GIDLNKDKFE KAMAVGATEC 240
 ISPKDSTKPI SEVLSEMTGN NVGYTFEVIK HLETMIDALA SCHMYGTSV VVGVPSPAKM 300
 LTYDPMMLFT GRTWKGCVFG GLKSRDVPK LVTEFLAKKF DLDQLITHVL PFKKISEGFE 360
 LLNSGQSIRT VLTF

Seq ID NO: 102 DNA sequence
 Nucleic Acid Accession #: NM_006783.1
 Coding sequence: 1..786

35
40
45
50

1 11 21 31 41 51
 | | | | |
 ATGGATTGGG GGAAGCTGCA CACTTTCATC GGGGGTGTC ACAAACACTC CACCAGCATC 60
 GGAAGAGTGT GGATCAGT CATCTTTATT TTCCGAGTCA TGATCCTAGT GGTGGCTGCC 120
 CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTGCTCTGCA ACACACTGCA ACCGGGATGC 180
 AAAAATGTGT GCTATGACCA CTTTTCCTCG GTGTCCACA TCCGGCTGTG GGCCCTCCAG 240
 CTGATCTTGG TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCCTA CTACAGGCAC 300
 GAAACCACTC GCAAGTTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAGA CATAGAGGAC 360
 ATTAAGAAAG ACAAGGTTTG GATAGAGGGG TCGCTGTGGT GGAAGTACAC CAGCAGCATC 420
 TTTTTCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTCTCT TTACAATGGG 480
 TACCACCTGC CCTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCAACCT GTTGTACTGC 540
 TTTATTTCTA GGCCCAACAGA GAAGACCGTG TTTACCATTT TTATGATTTC TCGCTCTGTG 600
 ATTTGCATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTITAGG 660
 AGATCAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720
 CAGAATGAAA TGAATGAGCT GATTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780
 AGCTAA

Seq ID NO: 103 Protein sequence:
 Protein Accession #: NP_006774.1

55
60

1 11 21 31 41 51
 | | | | |
 MDWGLTHTFI GGVNKHSTSI GKVVITVIFI FRVMILVVAA QEVWGDQED FVCNTLQPGC 60
 KNVICYDHFFP VSHIRLWALQ LIPVSTPALL VAMHVAYYRH ETTRKFRERG KRNDPKDIED 120
 IKKKVRIEG SLWWTYTSI PFRIIFEAAP MYVIFYLYNG YHLPWVLKCG IDPCPNLVDC 180
 FISRPTEKTV FTIFMISASV ICMILNVAEL CYLLKVCPR RSKRAQTQKN HPNHALKESK 240
 QNEMNELISD SGQNAITGPF S

Seq ID NO: 104 DNA sequence
 Nucleic Acid Accession #: NM_020411
 Coding sequence: 86-526

70
75
80

1 11 21 31 41 51
 | | | | |
 GGACCTGGGA AGGAGCATAG GACAGGGCAA GCGGGGATAA GGAGGGGCAC CACAGCCCTT 60
 AAGGCAGGAG GGAACCTCAC TGCGCATGCT CCTTTGGTGC CCACCTCAGT GCGCATGTTT 120
 ACTGGGCGTC TTCCCATCGG CCCCTTCGCC AGTGTGGGGA AGCGGGCGGA GCTGTGAGCC 180
 GGCAGCTCGG GTCCCTGAGG TCTGGATTCT TTCTCCGCTA CTGAGACAGG GCGGACACAC 240
 ACAAACACAG AACCAACAG CCACTCCAG GAGCCAGTA ATGGAGAGCC CCAAAAAGAA 300
 GAACCAAGCAG CTGAAAGTCC GGATCCTACA CCTGGGCGAG AGACAGAAGA AGATCAGGAT 360
 ACAGCTGAGA TCCCAAGTGC CGACATGGAA GGTGATCTGC AAGAGCTGCA TCAGTCAAAC 420
 ACCGGGGATA AATCTGGATT TGGGTTCGGG CGTCAAGGTG AAGATAATAC CTAAGAGGGA 480
 ACACGTGAAA ATGCCAAGAG CAGGTGAAGA GCAACCAAA GTTTAAATGA AGACAGCTGT 540
 AAACAACGCA AGCTGGTTTT ATATTAGATA TTTGACTTAA ACTATCTCAA TAAAGTTTTG 600
 CAGCTTTCAC CAAAAA AAAAA

Seq ID NO: 105 Protein sequence:
 Protein Accession #: NP_065144.1

85
1 11 21 31 41 51

MLLWCPPQCA CSLGVFPSP SPVWGTRRSC EPATRVPEVM ILSPLLRHGG HTQTQNHAS 60
PRSPVMESPK KKNQQLKVGI LHLGSRQKRI RIQLRSQCAT WKVICKSCIS QTPGINLDLG 120
SGVKVKIIPK EEHCMPPEAG EEQPVQ

Seq ID NO: 106 DNA sequence
Nucleic Acid Accession #: J04129
Coding sequence: 99-587

10 1 11 21 31 41 51
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCCTCTGC 60
TCACCCTGGG CGTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120
AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CCTGGCACTC CATGGCCATG GCGACCAACA 180
15 ACATCTCCCT CATGGGACAC CTGAAGGCCC CTCTGAGGGT CCACATCACC TCACTGTTGC 240
CCACCCCGGA GGACAACTGT GAGATCGTTC TGACAGATG GGAGAACAAC AGCTGTGTTG 300
AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG 360
TGGCGAAGCA GGCCACGCTG CTGATACTG ACTACGACAA TTCTCTGTTT CTCTGCTTAC 420
AGGACACCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCTGTGGTG 480
20 AGGACGATGA GATCATGCAG GGATTTCAT GGGCTTTTCA GCGCTTGGCC AGGCACCTAT 540
GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCGGTGCGG TTCTAGCTTC ACCTCGGCTC 600
CCAGGAAGAC CAGACTGCCA CCTTCCACA CCTCCAGAGC AGTGGGACTT CCTCTGCCCC 660
TTTCAAGAA TAACACAGC TCAGAAGACG ATGACGTGGT CATCTGTGTC GCCATCCCTT 720
25 TCCTGTGCA CACTGCACC ATTGCCATGG GGAGGCTGCT CCCTGGGGGC AGAGTCTCTG 780
GCAGAGGTTA TTAATAAAC CTTGGAGCAT G

Seq ID NO: 107 Protein sequence:
Protein Accession #: AAA60147

30 1 11 21 31 41 51
MDIPQTKQDL ELPLKAGTWH SMAMATNNIS LMATLKAPLR VHITSLPTP EDNLEIVLHR 60
WENNSCVEKK VLGEKTGNPK KFKINYTVAN EATLLDTDYD NFLFLCLQDT TPIQSMCMQ 120
35 YLARVLVEDD EIMQGFIRAF RPLPRHLWYL LDLKQMEEPK RF

Seq ID NO: 108 DNA sequence
Nucleic Acid Accession #: Bos sequence
Coding sequence: 48-794

40 1 11 21 31 41 51
TCCCAGGCAG CAGTTAGCCC GCGCCCGGCC TGTGTGTCCC CAGAGCCATG GAGAGAGCCA 60
GTCTGATCCA GAAGGCCAAG CTGGCAGAGC AGGCCGAAGC CTATGAGGAC ATGGCAGCCT 120
45 TCATGAAGAG CGCCGTGGAG AAGGGCGAGG AGCTCTCCTG CGAAGAGCGA AACCTGCTCT 180
CAGTAGCCTA TAAGAACTGT GTGGGCGGCC AGAGGGCTGC CTGGAGGGTG CTGTCCAGTA 240
TTGAGCAGAA AAGCAACGAG GAGGGCTCGG AGGAGAAGGG GCCCGAGGTG CGTGAGTACC 300
GGGAGAAGGT GGAGACTGAG CTCCAGGGCG TGTGCCACAC CGTGCTGGGC CTGCTGGACA 360
GCCACCTCAT CAAGGAGGCC GGGGACGCGG AGAGCGGGT CTCTACCTG AAGATGAAGG 420
50 GTGACTACTA CGCTACCTG GCGAGGTGG CCACCGGTGA CGACAAGAAG CGCATCATTG 480
ACTCAGCCCG GTGAGCCTAC CAGGAGGCCA TGGACATCAG CAAGAAGGAG ATGCCGCCCA 540
CCACCCCAT CGCCCTGGGC CTGGCCCTGA ACTTTTCCGT CTCTCACTAC GAGATCGCCA 600
ACAGCCCGGA GGAGGCCATC TCTCTGGCCA AGACCACCTT CGACGAGGCC ATGGCTGATC 660
TGCACACCTC CAGCGAGGAC TCCTACAAAG ACAGCACCTT CATATGCAG CTGCTGCGAG 720
55 ACAACCTGAC ACTGTGGAGC GCGGACAACG CCGGGGAAGA GGGGGGCGAG GCTCCCGAGG 780
AGCCCCAGAG CTGAGTGTGT CCGGCCACCG CCCCCTCCAG TCCCCACCCC 840
TCCGAGAGG ACTAGTATGG GGTGGGAGGC CCCACCTTC TCCCTAGGC GCTGTCTTGT 900
CTCCAAGGG CTCCGTGGAG AGGGACTGGC AGAGCTGAGG CCACCTGGGG CTGGGGATCC 960
60 CACTCTTCTT GCAGCTGTTG AGCGCACCTA ACCACTGTGT ATGCCCCAC CCCTGTCTCT 1020
CGCACCCGCT TCCCTCCGAC CCCAGGACCA GGCTACTTCT CCCCCTCTCT TGCCCTCCCT 1080
CTGCCCTGCG TGCCCTGATG CGTAGGAATT GAGGAGTGTG CCGCCTTGTG GCTGAGAACT 1140
GGACAGTGGC AGGGGCTGGA GATGGGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG 1200
CGCGCGCGCC AGTGCAAGAC CGAGATTGAG GGAAAGCATG TCTGCTGGGT GTGACCATGT 1260
65 TTCTCTCAA TAAAGTCCC CTGTGACACT C

Seq ID NO: 109 Protein sequence:
Protein Accession #: NP_006133.1

70 1 11 21 31 41 51
MERASLIQKA KLAEEAERYE DMAAFMKGAV EKGEELSCEE RNLLSVAYKN VVGQRAAWR 60
VLSSIEQKSN EESSEKQPE VREYREKVT ELQGVCDTVL GLDSDHLIKE AGDAESRVFY 120
LKMKGDIYRY LAEVATGDDK KRIIDSARSA YQEAIDISKK EMPPTNPIRL GLALNFSVPH 180
75 YEIANSPEEA ISLAKTTFDE AMADLHTLSE DSYKDSILIM QLRLDNLTLW TADNAGEEGG 240
EAPQEPQS

Seq ID NO: 110 DNA sequence
Nucleic Acid Accession #: NM_000695
Coding sequence: 407-1564

80 1 11 21 31 41 51
CACGAGTTGG TTTGGGAGCT GCCAGTCTCC TGGGAGGATC GCAGTCAGCA GAGCAGGGCT 60
GAGGCCTGGG GGTAGGAGCA GAGCCTGGCG ATCTGGAGGC AGCATGTCCA AGAAAGGGAG 120
85 TGGAGGTGCA GCGAAGGACC CAGGGGCGAG GCCACGCTG GGGATGGACC CCTTGGAGGA 180
CACACTGCGG CGGCTGCGTG AGGCCTTCAA CTGAGGGGCG ACGCGGCGCG CCGAGTTCCG 240
GGCTGCGCAG CTCACGGGCC TGGGCCACTT CCTTCAAGAA AACAAGCAGC TTCTGCGCGA 300

	CGTGCTGGCC	CAGGACCTGC	ATAAGCCAGC	TTTCGAGGCA	GACATATCTG	AGCTCATCCT	360
	TTGCCAGAAC	GAGGTGACT	ACGCTCTCAA	GAACCTTCAG	GCCTGGATGA	AGGATGAACC	420
	ACGGTCCACG	AACCTGTTC	TGAAGCTGGA	CTCGGTCTTC	ATCTGGAAGG	AACCCCTTGG	480
5	CCTGGTCTCT	ATCATCGCAC	CCTGGAACTA	CCCATTTGAAC	CTGACCCCTG	TGCTCCTGGT	540
	GGGCAACCTC	CCCGCAGGGA	ATTGGTGGT	GCTGAAGCCG	TCAGAAATCA	GCCAGGGCAC	600
	AGAGAAGGTC	CTGGCTGAGG	TGCTGCCCCA	GTACCTGGAC	CAGAGCTGCT	TTGCCGTGGT	660
	GCTGGGGGGA	CCCGAGGAGA	CAGGGCAGCT	GCTAGAGCAC	AAGTTGGACT	ACATCTCTTT	720
	CACAGGGAGC	CCTCGTGTGG	GCAAGATTGT	CATGACTGCT	GCCACCAAGC	ACCTGACGCC	780
10	TGTCAACCTG	GAGCTGGGGG	GCAAGAACCC	CTGCTACGTG	GACGCAACT	GCGACCCCA	840
	GACCGTGGCC	AACCGGTGG	CCTGGTCTG	CTACTTCAAT	GCCGGCCAGA	CCTGCGTGGC	900
	CCCTGACTAC	GTCTGTGTGA	GCCCGAGAT	GCAGGAGAGG	CTGCTGCCCG	CCCTGCAGAG	960
	CACCATCACC	CGTTTCTATG	GCGACGACCC	CCAGAGCTCC	CCAAACCTGG	GCGCATCAT	1020
	CAACCAGAAA	CAGTTCACG	GGCTGCGGGC	ATTGCTGGGC	TGCGGCGCGC	TGGCCATTGG	1080
15	GGGCAAGAGC	AACGAGAGCG	ATCGCTACAT	CGCCCCACG	GTGCTGTTGG	ACGTGCAGGA	1140
	GACCGAGCCT	GTGATGCAGG	AGGAGATCTT	CGGGCCCATC	CTGCCCATCG	TGAACGTGCA	1200
	GAGCGTGGAG	GAGGCCATCA	AGTTTCATCA	CCGGCAGGAG	AAGCCCTTGG	CCCTGTACGC	1260
	CTTCTCCAAC	AGCAGACAGG	TTGTGAACCA	GATGCTGGAG	CGGACCAGCA	GCGGCAGCTT	1320
	TGGAGGCAAT	GAGGGCTTCA	CCTACATATC	TCTGCTGTCC	GTGCCATTCC	GGGGAGTCCG	1380
20	CCACAGTGGG	ATGGGCCGGT	ACCACGGCAA	GTTCACCTTC	GACACCTTCT	CCCAACACCG	1440
	CACCTGCCCTG	CTCGCCCTG	CGGGCCTGGA	GAAATTAAAG	GAGATCCGCT	ACCCACCTTA	1500
	TACCGACTGG	AACCAAGCAG	TGTTACGCTG	GGGCATGGGC	TCCAGAGCT	GCACCTCTCT	1560
	GTGAGCGTCC	CACCGCCTC	CAACGGGTCA	CACAGAGAAA	CCTGAGTCTA	GCCATGAGGG	1620
	GCTTATGCTC	CCAACCTACA	TTGTTCTCTC	AGACCGCAGG	CTCCCCCAGC	CTCAGSTTGC	1680
25	TGGAGCTGTC	ACATGACTGC	ATCCTGCCTG	CCAGGGCTGC	AAAGCAAGGT	CTTGCTTCTA	1740
	TCTGGGGGAC	GCTGCTCGAG	AGAGGCCGAG	AGGCCGAGA	ACATGCCAGT	TGCTCTCACT	1800
	CACCCCAACC	TCCCAATTC	CAGCCCTTGG	CCCTCTCGGT	CAGGGTTGGC	CAGGCCCACT	1860
	CACAGGGGCA	GTGTCACTCC	GGAAAATACA	GTGCCCTGCC	TTCTTAGGGG	CATCAGCCCT	1920
	GACGGTGTGA	GAGCGTGGAG	CCCTCCAGGC	CTTGTCTCTC	CCCTCTAGGC	ACACGGGCAC	1980
30	TTCCACCTCT	GCCTCATCCC	AACTGCACCA	GCACCTGCTC	CCCAAGGAT	CCTCTCACAT	2040
	CCCACTCTG	TCTCTGCACC	ACCCCTCTGG	TTACACCGC	ACCTGCACT	CACCCACAGC	2100
	AGCTCCATCC	ACTGGGAAAA	CTGGGGTTTG	CATCACTCCA	CTGCACAGTG	TTAGTGGGAC	2160
	CTGGGGGCAA	GTCCCTTGAC	TTCTCTGAGC	CTCAGTTTCC	TTATGTGAAA	GTTGCTGGAA	2220
	CCAAAATGGA	GTCACTTATG	CCAACTCTA	ATAAATGGA	GTGGGGGGG	CACATAGAAG	2280
35	CCCTCACACA	CACATGCCCC	TAACAGGATT	TATCACCAAG	ACAAGCCTGC	ATGTAAGACC	2340
	AGACACAGGG	GTCTGTGAAA	AGCAGTCTCT	CAAAAGACTG	AGTATTCCAG	ATGAGCTGCA	2400
	GATGCTTACC	TACCAAGGCC	GTCTCCACCA	GAAAACCATC	GCCCACTCCT	GOGATCAGCT	2460
	TGTGACTTAC	AAACCTTGTT	TAAAAGCTGC	TTACATGGAC	TTCTGTCTCT	TAAAACGTTT	2520
40	CCCTTGGCTG	TGGCCCTCTG	TGTATGCCTG	GGATCCTTCC	AAGCACTCAT	AGCCACAGATA	2580
	GGAACTCTCT	GCTCTCTCCA	AATAAATTCA	TCTGTTCT			

Seq ID NO: 111 Protein sequence:
Protein Accession #: NP_000686

45	1	11	21	31	41	51	
	MDDEPRSTNL	FMKLDVSFVW	KEPFGLVLII	APWNYPLNLT	LVLLVGLTLP	GNCVVLKPSE	60
	ISQGTKEVLA	EVLQPYLDQS	CPAVVLGGPQ	ETGQLLEHL	DYIFPTGSPR	VGKIVMTAAT	120
50	KHLTPVTLEL	GGKNPCYVDD	NCDPQTVANR	VAWPCYFNAG	QTCVAPDYVL	CSPEMQERLL	180
	PALQSTITRF	YDDDPQSSFN	LGRIINQKQF	QRLRLALLGC	RVAIGGQSNB	SDRYIAPTVL	240
	VDVQETPEVM	QSEIFPGPILP	IVNVQSVDEA	IKFINRQKEP	LALYAFNSNR	QVNVQMLERT	300
	SSGSPGQNEG	FTYISLLSVP	PGGVGHSOMG	RHYHGKFTPD	FSHRTCLLA	PSGLEKLKEI	360
	RYPPYTDWQ	QLLRWGMGSQ	SCTLL				

Seq ID NO: 112 DNA sequence
Nucleic Acid Accession #: NM_004456
Coding sequence: 58-2298

60	1	11	21	31	41	51	
	GAATTCGGGG	CGACGCGCGG	GAACAACGCG	AGTCGGCGCG	CGGGAAGAG	AATAATCATG	60
	GGCCAGACTG	GGAGAAGATC	TGAGAAGGGA	CCAGTTTGT	GGCGAAGCG	TGTAAATCA	120
	GAGTACATGC	GACTGAGACA	GCTCAAGAGG	TTCAAGAGG	CTGATGAAGT	AAAGAGTATG	180
65	TTTAGTTCCA	ATCGTCAGAA	AATTTTGAA	AGAACGGAAA	TCTTAAACCA	AGAATGGAAA	240
	CAGCGAAGGA	TACAGCCTGT	GCACATCCTG	ACTTCTGTGA	GCTCAATTGG	CGGGACTAGG	300
	GAGTGTTCGG	TGACAGTGA	CTTGATTTT	CCAACACAG	TCATCCCAT	AAAGACTCTG	360
	AATGCAGTTG	CTTCAGTACC	CATAATGTAT	TCTTGGTCTC	CCCTACAGCA	GAATTTTATG	420
	GTGGAAGATG	AAACTGTITT	ACATAACATT	CCTTATATGG	GAGATGAAGT	TTTAGATCAG	480
70	GATGTAATCT	TCAATGAAGA	ACTAATAAAA	AATTATGATG	GGAAAGTACA	CGGGATAGA	540
	GAATGTGGGT	TTATAAATGA	TGAAATTTT	GTGGAGTTGG	TGAATGCCCT	TGGTCAATAT	600
	AATGATGATG	ACGATGATGA	TGATGGAGAC	GATCCTGAAG	AAAGAGAAGA	AAAGCAGAAA	660
	GATCTGGAGG	ATCACCGAGA	TGATAAAGAA	AGCCGCCAC	CTCGAAATT	TCCTTCTGAT	720
	AAAAATTTTG	AGGCCATTTC	CTCAATGTTT	CCAGATAAGG	GCACAGCAGA	AGAACTAAG	780
75	GAAAAATATA	AAGAACTCAC	CGAACAGCAG	CTCCAGGCG	CACCTCTCTC	TGAATGTACC	840
	CCCAACTAG	ATGGACCAAA	TGCTAAATCT	GTTCAGAGAG	AGCAAAGCTT	ACACTCTTTT	900
	CATACGCTTT	TCTGTAGGCG	ATGTTTAAA	TATGACTGCT	TCTACATCC	TTTTCATGCA	960
	ACACCCAAACA	CTTATAAGCG	GAAGAACACA	GAAACAGCTC	TAGACAAACA	ACCTTGTGGA	1020
	ACACAGTTGT	ACCAAGATT	GGAGGGAGCA	AAGGAGTTTG	CTGCTGCTCT	CACCGCTAG	1080
80	CGGATAAAGA	CCCCACCAA	ACGTCCAGGA	GGCCGAGAA	GAGGACGGCT	TCCCAATAAC	1140
	AGTAGCAGGC	CCAGCACCCC	CACCATTAAT	GTGCTGGAAT	CAAAGGATAC	AGACAGTGAT	1200
	AGGGAAGCAG	GGACTGAAC	GGGGGAGAG	AACAATGATA	AAGAAGAAGA	AGAGAAGAAA	1260
	GATGAACACT	CGAGCTCTCT	TGAAGCAAA	TCTCGTGTCT	AAACACCAAT	AAAGATGAAG	1320
	CCAAATATTG	AACCTCTCTG	GAATGTGGAG	TGGAGTGGTG	CTGAAGCCTC	AATGTTTGA	1380
85	GTCTCTATTG	GCACCTACTA	TGACAATTC	TGTGCCATTG	CTAGGTTAAT	TGGGACCAAA	1440
	ACATGTAGAC	AGGTGATGGA	GTTTAGAGTC	AAAGAATCTA	GCATCATAGC	TCCAGCTCCC	1500
	GCTGAGGATG	TGGATACTCC	TCCAAGGAAA	AAGAAGAGGA	AACACCGGTT	GTGGCTGCA	1560
	CATGTCAGAA	AGATACAGCT	GAAAAGGAC	GGCTCTCTA	ACCATGTTTA	CAACTATCAA	1620

	CCCTGTGATC	ATCCACGGCA	GCCTTGTGAC	AGTTCGTGCC	CTTGTGTGAT	AGCACAAAAT	1680
	TTTTTGTAAA	AGTTTTGTCA	ATGTAGTTCA	GAGTGTCAAA	ACCGCTTTCC	GGGATGCCGC	1740
	TGCAAGACAC	AGTGCAACAC	CAAGCAGTGC	CCGTGCTACC	TGGCTGTCCG	AGAGTGTGAC	1800
5	CCTGACCTCT	GTCTTACTTG	TGGAGCCGCT	GACCAITGGG	ACAGTAAAAA	TGTGTCTCTG	1860
	AAGAACTGCA	GTATTACAGC	GGGCTCCAAA	AAGCATCTAT	TGCTGGCACC	ATCTGACGTG	1920
	GCAGGCTGGG	GGATTTTTAT	CAAAGATCCT	GTGCAGAAAA	ATGAATTCAT	CTCAGAATAC	1980
	TGTGGAGAGA	TTATTTCTCA	AGATGAAGCT	GACAGAAGAG	GGAAAGTGTA	TGATAAATAC	2040
	ATGTGCAGCT	TTCTGTTCAC	CTTGAACAAT	GATTTTGTGG	TGGATGCAAC	CCGCAAGGGT	2100
10	AACAAAATTC	GTTTTGTCAA	TCATTGCGTA	AATCCAAACT	GCTATGCAAA	AGTTATGATG	2160
	GTTAACGGTG	ATCACAGSAT	AGGTATTTTT	GCCAAGAGAG	CCATCCAGAC	TGGCGAAGAG	2220
	CTGTTTGTGG	ATTACAGATA	CAGCCAGGCT	GATGCCCTGA	AGTATGTCCG	CATCGAAAGA	2280
	GAAATGGAAA	TCCCTTGACA	TCTGCTACCT	CCTCCCCCTC	CTCTGAAACA	GCTGCCCTTAG	2340
	CTTCAGGAAC	CTCGAGTAGT	GTGGGCAATT	TAGAAAAAGA	ACATGCAGTT	TGAAATTCTG	2400
15	AAATTGCAAA	TACTGTGAAC	AATAATTTAT	AGTAATGAGT	TTAAAAATCA	ACTTTTTATT	2460
	GCCTTCTCAC	CAGCTGCAAA	GTGTTTTGTA	CCAGTGAATT	TTTGCAATAA	TGCAGTATGG	2520
	TACATTTTTC	AACTTTGAAT	AAAGAATACT	TGAACTTGAA	AAAAAAAAAA	AAAAAA	

20 Seq ID NO: 113 Protein sequence:
Protein Accession #: NP_004447

	1	11	21	31	41	51	
25	MGQTGKKSEK	GPVCWRKRVK	SEYMRLRQLK	RPFRRADEVKS	MFSSNRQKIL	ERTBILNQEW	60
	KQRRIQPVHI	LTSVSSLRGT	RECSVTSDDL	FPTQVPLKLT	LNAVASVPIM	YWSPLQONF	120
	MVEDETVLIN	IPYMGDEVID	QDGTFFIEELI	KNYDGKVHGD	RECGFINDEI	FVELVNALGQ	180
	YNDDDDDDGG	DDPEEREKQ	KDLEDHRDDK	ESRPPRKFPF	DKILEAISSM	FPDKGTAEEL	240
	KEKYKELTQ	QLPGALPPEC	TPNIDGPNK	SVQREQSLHS	PHTLFCRRCF	KYDCFLHPPH	300
	ATPNTYKRVK	TETALDNKPC	GPQCYQHLBG	AKEPFAALTA	ERIKTPPKRP	GRRRRRLPN	360
30	NSSRPSTPTI	NVLESKDTDS	DREAGTETGG	ENNDKEEEKK	KDETSSSSEA	NSRCQTPIRM	420
	KPNIEPPENV	EWGGAESAMP	RVLIGTYIDN	FCATLARLIGT	KTCRQVYEFR	VKESIIAPA	480
	PAEDVDTPPR	KKGRKHLWA	AHCRKIQLKK	DGSSNHVYNY	QPCDHPROP	DSSCPCVIAQ	540
	NFCEKPCQCS	SEQCNRPFGC	RCKAQCNTRQ	CPCYLAVREC	DFDLCLTCGA	ADHWDKSNVS	600
35	CKNCSIQRGS	KKHLLLAPSD	VAGWGIFIKD	PVQKNEPISE	YCGEIIISQDE	ADRRGKVYDK	660
	YMCSEFLNIN	NDPVVDATRK	GNKIRPANHS	VNPNCYAKVM	MVNGDHRIGI	FAKRAIQTGE	720
	ELFVDYRYSQ	ADALKVVGIE	REMEIP				

40 Seq ID NO: 114 DNA sequence
Nucleic Acid Accession #: NM_001827
Coding sequence: 96-335

	1	11	21	31	41	51	
45	AGTCTCOGGC	GAGTGTGTGC	CTGGGCTGGA	CGTGGTTTTG	TCTGCTGCGC	COGCTCTTCG	60
	CGCTCTCGTT	TCATTTTCTG	CAGCGCGCCA	CGAGGATGGC	CCACAAGCAG	ATCTACTACT	120
	CGGACAAGTA	CTTCGACGAA	CACCTACGAGT	ACCGGCATGT	TATGTTACCC	AGAGAACTTT	180
	CCAAACAAGT	ACCTAAACT	CATCTGATGT	CTGAAGAGGA	GTTGAGGAGA	CTTGGTGTCC	240
	AACAGAGTCT	AGGCTGGGTT	CATTACATGA	TTCATGAGCC	AGAACCACAT	ATTCTTCTCT	300
50	TTAGAGGACC	TCTTCCAAAA	GATCAACAAA	AATGAAGTTT	ATCTGGGGAT	CGTCAAACTC	360
	TTTTCAAAAT	TAATGTATAT	GTGTATATAA	GGTAGTATTC	AGTGAATACT	TGAGAAATGT	420
	ACAAATCTTT	CATCCATACC	TGTGCATGAG	CTGTATTCTT	CACAGCAACA	GAGCTCAGTT	480
	AAATGCACCT	GCAAGTAGTG	TACTGTGAAGA	TGTTTAAGAT	AAAAGTTCTT	CCAGTCAGTT	540
	TTTCTCTTAA	GTGCGCTGTT	GAGTTTACTG	AAACAGTTTA	CTTTTGTTC	ATAAAGTTTG	600
55	TATGTTGCAT	TTAAAAAAA	AAAAAAA				

Seq ID NO: 115 Protein sequence:
Protein Accession #: NP_001818

	1	11	21	31	41	51	
60	MAHKQIYSD	KYFDEHYEYR	HVMLPRELSK	QVPKTHLMSE	EEWRLRGVQQ	SLGWVHYMIH	60
	EPEPHILLPR	RPLPKDQKQ					

65 Seq ID NO: 116 DNA sequence
Nucleic Acid Accession #: CAT cluster

	1	11	21	31	41	51	
70	TCAGACCTCA	TGAGTCACTT	GGACTCTTGA	GCCACCTCTG	GGGGTGGAGT	CTCTCTCCTG	60
	GCATCTGGAC	CCTTGGTGCT	ATCGACGAAG	CTTGGGTGGG	GCTCTTAGCT	GCTATGTGCA	120
	AGAGGTGTGT	TCCAGGGAAG	GCCCCTATCT	CTCTGCAGAG	GTCAAGTGAA	AGCGACGGCC	180
	GCAGCCAACA	GAGTTCAAAA	TGCAGGCTTG	GAAAGTACAG	GGGGCTCTGT	GGAGGATGGG	240
	AAGGACTGAT	CCACATTCCC	ACCAGGAAGT	TTAGCAGAAC	CCCCGCGTGC	CAACTGGACC	300
75	CCTTGGAAAG	ACCTGGCTCA	GGCTGGACCA	CCTCTTGAGA	GGGAGGAGCT	CTGGATTGGA	360
	TCAGAAATTC	TTTGTGAGC	ATGGTGCTTC	ATGCTATATA	TACCAACACT	TTGGGAGGCC	420
	AGTGTGGGAG	GATCTCTTGA	GCCAGGAGT	TCAAGACTAG	CCTGGGCAAC	ACAGAGAGAA	480
	CCCATCTCTA	AAATAATAAT	AATAATAAAA	TAAAAAATTA	GCAGGGCATG	GTGGCATGTG	540
	CCTGTAGTTC	CAGCTACCCA	GGAGGCTGAG	GCAAGAGGAT	GGCTGGAGCC	TGGGATGTTG	600
	AGGCTGCAAT	GAACTGTGAT	TACCCCACTG	CACCTCCAGC	TGGGCAAAAG	AGCGAGAGAA	660
80	CTGTCTCTCA	ATAATAATAA	TAATAATAAT	CTTATTTTGG	AGAATAAAGA	GACCTCTGGA	720
	TTTGAAGTGC	CATTTGGGTA	GAAAGAAAAG	ACGITTACAC	CGAGAAATAG	TCTGTGTTGC	780
	CCTGAAGGAG	CAGAGGGATG	CATGCTGGGA	GGTGACCTAC	AGTTGAAGAA	GACTCATTAT	840
	GACAGACCTT	GTCCCTCTTC	CTTGTGGAAA	GTGTTTCTTC	TGCTGCTACT	GCTCATGAGA	900
	CTCTTCCCCC	TCCCTGTCCC	AGGGAAACAA	AGGGCTTTCT	ACCAACCCCT	TTCTTGCCCC	960
85	CCGCTCTCCA	TGTCTGCTGT	GCCTTGTGAC	TCAGCAATTC	TTGTTTGCTC	CATTATCTTC	1020
	CAGCCGGATA	CAGAGTGAAT	AGTTAACAC	ACTTAGGTCA	AATAGGATCT	AAATTTTGT	1080
	TCTGTCTCCG	TGTAAGAGAG	CCAGTGTGTTG	TGTGTTGCAA	GCAGCCTTGG	AATAGTAACT	1140

WO 02/086443

PCT/US02/12476

5
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Seq ID NO: 117 DNA sequence
Nucleic Acid Accession #: BC012178.1
Coding sequence: 204-2285

1 11 21 31 41 51
CTTCTCTCCC GCGCGCTGG GCGCGGCT CCGCTGCTGT TGCTCCATTC GCGGCTTTTC 60
TGGCGGCTGG CTCTCTCCG CTGCGGCTG CTCTCGACC AGGCTCCTT CTCAACCTCA 120
GCGCGGCTGG CCGACCTTC CCGCACCTTC CCGCCCGTC TCGTACTGTC GCGCTCACCG 180
CGCGGCTGG CCGCTGCTG CCGATGCTG TGTGCAACG AGACTCCAAG CTGGAGAAATG 240
CTGGAGGAGA CCTTAAGAGT GCGCACCACT ACTATGAAG AGCTGTGTC ATTCTGGATG 300
CTGGTGTCTA GTACGGGAAA GTCATAGACC GAAGAGTGAG GGAAGTGTTC GTGCAGTCTG 360
AAATTTTCCC CTGGAACA CAAGCATTTG CTATAAGGA ACAAGGATTC CGTGTATTA 420
TCATCTCTGG AGGACCTAAT TCTGTGTATG CTGAAGATGC TCCCTGGTTT GATCCAGCAA 480
TATTACTAT TGGCAAGCCT GTTCTTGGAA TTGCTATGG TATGCAGATG ATGAATAAGG 540
TATTGGAGG TACTGTGCAC AAAAAAAGTG TCAGAGAAGA TGGAGTTTC AACATTAGTG 600
TGGATAATAC ATGTTTCTTA TTCAGGGGCC TTCAGAAGGA AGAAGTGTTC TTGCTTACAC 660
ATGAGATAG TGTAGACAAA GTAGCTGATG GATTCAAGGT TGTGCGACGT TCTGGAAACA 720
TAGTAGCAGG CATAGCAAAAT GAATCTAAAA AGTTATATGG AGCAGAGTTC CACCCTGAAG 780
TTGGCCTTAC AGAAAAATGGA AAGTAATAC TGAAGAATTT CCTTTATGAT ATAGCTGGAT 840
GCAGTGGAAC CTTACCGCTG CAGAACAGAG AACTTGAGTG TATTGAGAG ATCAAAGAGA 900
GAGTAGGCAC GTCAAAAGTT TTGGTTTAC TCAGTGGTGG AGTAGACTCA ACAGTTTGTA 960
CAGCTTTGCT AAATCGTGCT TTGAACCAAG AACAGTCAT TGCTGTGCAC ATTGATAATG 1020
GCCTTATGAG AAAACGAGAA AGCCAGTCTG TTGAAGAGGC CCTCAAAAAG CTGGAAATTC 1080
AGGTCAAAGT GATAAATGCT GCTCAATCTT TCTACAATGG AACACAACCT CTACCAATAT 1140
CAGATGAAGA TAGAACCCCA CGGAAAAGAA TTAGCAAAAC GTTAAATATG ACCCAAGTC 1200
CTGAAGAGA AAGAAAAATC ATTGGGGATA CTTTGTGTAA GATTGCCAAT GAAGTAATTG 1260
GAGAAATGAA CTGAAACCA GAGGAGGTTT TCCTTGCCCA AGGTACTTTA CGGCTGATC 1320
TAATTGAAG TGCATCCCTT GTTGCAAGTG GCAAGCTGA ACTCATCAA ACCCATCACA 1380
ATGACACAGA GCTCATCAGA AAGTTGAGAG AGGAGGGAAA AGTAATAGAA CCTCTGAAG 1440
ATTTTCATAA AGATGAAGTG AGAATTTTGG CGACAGAACT TGGACTTCCA GAAGAGTTAG 1500
TTTCCAGGCA TCCATTTCCA GGTCTGGGCC TGGCAATCAG AGTAATATGT GCTGAAGAAC 1560
CTTATATTG TAAGGACTTT CCTGAAACCA ACAATATTTT GAAATAGTA GCTGATTTT 1620
CTGCAAGTGT TAAAGGCCA CATACCTTAT TACAGAGAGT CAAAGCTGTC ACAACAGAA 1680
AGGATCAGGA GAGCTGATG CAAATTACCA GTCTGCATTC ACTGAATGCC TTCTGTGCTG 1740
CAATTAAAC TGTAGGTGTG CAGGTGACT GTCTGTCTCA CAGTTACGTG TGTGGAATCT 1800
CCAGTAAGA TGAACCTGAC TGGGAATCAC TTATTTTCTT GGCTAGGCTT ATACCTGCA 1860
TGTGTACAAA CGTTAACAGA GTTGTTTATA TATTGGGCC ACCAGTTAAA GAACCTCCTA 1920
CAGATGTTAC TCCCACTTTC TTGACAACAG GGGTGCTCAG TACTTTACGC CAAGCTGATT 1980
TTGAGGCCCA TAACATTTCT AGGGAGTCTG GGTATGCTGG GAAATCAGC CAGATGCCGG 2040
TGATTTGAC ACCATTACAT TTTGATCGGG ACCCACTTCA AAAGCAGCCT TCATGCCAGA 2100
GATCTGTGTT TATTGCAACC TTTATTACTA GTGACTTCAT GACTGTATA CCGCAACAC 2160
CTGGCAATGA GATCCCTGTA GAGGTGTTAT TAAAGATGGT CACTGAGATT AAGAAGATT 2220
CTGATTTTTC TCGAATTATG TATGACTTAA CATCAAAGCC CCCAGGAAT ACTGAGTGG 2280
AGTAATAAAC TTCTGTCTCT ATTAATA

60 Seq ID NO: 118 Protein sequence:
Protein Accession #: AAL12178.1

1 11 21 31 41 51
MALCNGDSKL ENAGGDLKDG HHHYEGAVVI LDAGAQQYKIV IDRRVRELIV QSEIFPLETP 60
APAIKQDQFR AIIISGGPNS VYAEADAPWFD PAIFTIGKPV LGICYGMQMM NKVPGTVHK 120
KSVREDGVFN ISVDNTCSLF RGLQKEBVVL LTHGDSVDKV ADGFKVARS GNIVAGIANE 180
SKKLYGAQFH PEVLGTENGK VILKNFLYDI AGCSGTPTVQ NRELECIREI KERVGTSKVL 240
VLLSGGVDSV VCTALLNRL NQEQVIAVHI DNGFMRKRES QSVVEALKKL GIQVKVINAA 300
HSFYNGTTTL PISDEDRTPR KRISKTLMNT TSPEKRKII GDTFVKIANE VIGEMNLKPE 360
EVFLAQGLTL PDLIESASLV ASGKAELIKT HENDTELIRK LREEGKVIK LKDPHKDEV 420
ILGRELGPE ELVSRHPPPG PGLAIRVICA EEPYICKDFF ETNNILKIVA DPSASVKKPH 480
TLLQVRKACT TEEDQEKLMQ ITSLSLSLNAF LLPIKTVGVQ GDCRSYSVVC GISSKDEPDW 540
ESLIFLARLI PRMCHNVNRV VYIFGPPVKE PPTDVTPFTL TTGVLSTLRQ ADPEAHNLR 600
ESGYAGKISQ MPVILTPLHF DRDPLQKQPS QORSVIRTIF ITSDPMTGIP ATPGNEIPVE 660
VVLQWVTEIK KIPGISRIMY DLTSKPPGTT EWE

80 Seq ID NO: 119 DNA sequence
Nucleic Acid Accession #: NM_006500.1
Coding sequence: 27..1967

1 11 21 31 41 51
ACTTGCCTCT CGCCCTCOGG CCAAGCATGG GCGTTCCAG GCTGGTCTGC GCCTTCTTGC 60
TCGCGCCTG CTCTGCTGCT CCTCGCTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
CGCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCTT TCTGAAGTGC GGCCTCTCCC 180
AGTCCCAAG CCACTCAGC CATGTCGACT GGTTCCTGT CCACAAGGAG AAGCGGACGC 240

	TCATCTTCGG	TGTGCGCCAG	GGCCAGGGCC	AGAGCGAACC	TGGGAGTAC	GAGCAGCGGC	300
	TCAGCCTCCA	GGACAGAGGG	GCTACTCTGG	CCCTGACTCA	AGTCACCCCC	CAAGACGAGC	360
	GCATCTTCTT	GTGCCAGGGC	AAGCGCCCTC	GGTCCCAGGA	GTACCGCATC	CAGCTCCGCG	420
5	TCTACAAAGC	TCCGAGAGAG	CCAAACATCC	AGGTCAACCC	CCTGGGCATC	CCTGTGAACA	480
	GTAAAGAGCC	TGAGGAGGTC	GCTACCTGTG	TAGGAGAGAA	CGGTACCCCC	ATTCTCTCAAG	540
	TCATCTGGTA	CAAGAAATGGC	CGGCTCTGTA	AGGAGGAGAA	GAACCGGGTC	CACATTCAGT	600
	CGTCCCAGAC	TGTGGAGTCG	AGTGGTTTGT	ACACCTTGCA	GAGTATTCTG	AAGGCACAGC	660
	TGGTTAAAGA	AGACAAGAT	GCCCAGTTTT	ACTGTGAGCT	CAACTACCGG	CTGCCCAGTG	720
10	GGAAACCAAT	GAAGGAGTCC	AGGGAAGTCA	CCGTCCCTGT	TTTCTACCCG	ACAGAAAAAG	780
	TGTGGCTGGA	AGTGGAGCCC	GTGGGAATGC	TGAAGGAAGG	GGACCGCGTG	GAATCAGGTT	840
	GTTTGGCTGA	TGGCAACCCCT	CCACCAACAT	TCAGCATCAG	CAAGCAGAAC	CCCAGCACCA	900
	GGGAGGCAGA	GGAAAGAGACA	ACCAACGACA	ACGGGGTCTT	GGTGTGAG	CCTGCCCGGA	960
	AGGAACACAG	TGGGCGCTAT	GAATGTCAAG	CCTGGAACCT	GGACACCATG	ATATCGCTGC	1020
	TGAGTGAACC	GAAGGAACTA	CTGGTGAAC	ATGTGTCTGA	CGTCCGAGTG	AGTCCCGCAG	1080
15	CCCTCAGAG	ACAGGAAGCG	AGCAGCCTCA	CCCTGACCTG	TGAGGCAGAG	AGTAGCCAGG	1140
	ACCTCGAGTT	CCAGTGGCTG	AGAGAAGAGA	CAGACCAGGT	GCTGGAAGG	GGGCTGTGTC	1200
	TTCACTTGCA	TGACCTGAAA	CGGGAGGCAG	GAGGCGGCTA	TGCTGCGTG	CGCTCTGTGC	1260
	CCAGCATACC	CGGCGTGAA	CGCACACAGC	TGGTCAAGCT	GGCCATTTTT	GGCCCCCTTT	1320
20	GGATGGCATT	CAGGAGAGAG	AAGGTGTGGG	TGAAGAGAGAA	TATGGTGTGG	AATCTGTCTT	1380
	GTGAAGCGTC	AGGGCACCCC	CGGCCACCA	TCTCTCTGAA	CGTCAACGCG	ACGGCAAGTG	1440
	AACAAGACCA	AGATCCACAG	CGAGTCTCTG	GCACCTTGAA	TGTCTCTGTG	ACCCCGGAGC	1500
	TGTTGGAGAG	AGGTGTTGAA	TGCACGGCCT	CCACCGACCT	GGGCAAAAAC	ACCAGCATCC	1560
	TCTTCTCTGA	GCTGGTCAAT	TTAACCAACC	TCACACACAG	CTCCAAACACA	ACCCTGGGCC	1620
25	TCAGCACTTC	CAGTGGCAGT	CCTCATACCA	GAGCCAAACAG	CACCTCCACA	GAGAGAAAGC	1680
	TGCCGAGGCC	GGAGAGCCGG	GGCGTGGTCA	TGCTGGCTGT	GATTGTGTGC	ATCTGTGTCC	1740
	TGGCGGTGCT	GGGCGCTGTC	CTCTATTTC	TCTATAAGAA	GGGCAAGCTG	CCGTGCAGGC	1800
	GCTCAGGAA	CAGGAGATC	ACGCTGCCCC	CGTCTCGTAA	GACCGAACTT	GTAGTTGAAG	1860
	TTAAGTCAGA	TAAAGTCCCA	GAAGAGATGG	GCCTCCTGCA	GGGCAAGCAG	GGTGACAAAG	1920
30	GGGCTCCGGG	CAGTGGCAGT	GAGAAATACA	TGATCTGAG	GCATTGACCC	CGAATCACTT	1980
	CAGCTCCCTT	CCCTGCCCTG	ACCATTCCCA	GCTCCCTGCT	CACCTCTCTC	TCAGCCAAAG	2040
	CCTCCAAAGG	GACTAGAGAG	AAGCCTCCTG	CTCCCTCAC	CTGCACACCC	CCTTTCAGAG	2100
	GGCCACTGGG	TTAGGACCTG	AGGACCTCAC	TTGGCCCTGC	AAGCCGCTTT	TCAGGGACCA	2160
	GTCCACACCC	ATCTCTCTCA	CGTTGAGTGA	AGCTCATCCC	AAGCAAGGAG	CCCCAGTCTC	2220
35	CGAGCGGGT	AGGAGGTTT	CTTGCAAGAC	GTGTTTTTTC	TTTACACACA	TTATGGCTGT	2280
	AAATACCTGG	CTCCTGCCAG	CAGCTGAGCT	GGGTAGCCTC	TCTGAGCTGG	TTTCTGCCCC	2340
	CAGAGGCTGG	CTTCCACCAT	CCAGGTGCAC	CAGTGAAGTG	AGGACACACC	GGAGCCAGGC	2400
	GCCTGCTCAT	GTGAAGTGG	GCTGTTTACA	CCCGCTCCGG	AGAGCACCCC	AGCGGCATCC	2460
	AGAAGCAGCT	CGAGTGTTCG	TGCCACACCC	CTCCTGCTCG	CCTCTTCAAA	GTCTCCTGTG	2520
40	ACATTTTTTT	TTTGGTTCGA	AGCCAGGAAC	TGGTGTCTAT	CCTTAAAGA	TACGTGCGGG	2580
	GGCCAGGTGT	GGTGGCTCAC	GCCTGTAATC	CCAGCACTTT	GGGAGGCCGA	GGCGGGCGGA	2640
	TCACAAAGTC	AGGACGAGAC	CATCCTGGCT	AACACGGTGA	AACCTGTGCT	CTACTAAAAA	2700
	TACAAAAAAA	AATTAGCTAG	CGGTAGTGGT	TGGCACCTAT	AGTCCAGCT	ACTCGGAAGG	2760
	CTGAAGCAGG	AAGATGTTAT	GAATCCAGGA	GGTGAGGCTT	GCAGTGAGCC	GAGACCGTGC	2820
45	CAGTCACTC	CAGCTGGGG	AACACAGCGA	GACTCCGCTC	CGAGGAAAAA	AAAAGAAAAA	2880
	ACGGTACCT	CGGGTGAAGA	AGCTGGGCGC	TGTTTTCGAG	TTGAGGTGAA	TTAGCCTCAA	2940
	TCCCGGTGTT	CACCTGCTCC	CATAGCCCTC	TTGATGGATC	ACGTAAAACT	GAAAGGCAGC	3000
	GGGAGCAGCA	CAAGAGTAGG	GTCTACACTG	TCCTTCATGG	GGATTAAGC	TATGGTTATA	3060
	TTAGCACCAA	ACTTCTACAA	ACCAAGCTCA	GGGCCCAAC	CCTAGAAAGG	CCCAATAGAG	3120
50	AGAATGGTAT	TAGGGATAG	AAAAAGGGGC	CTGGCTAGAG	CTTCCGGTGT	GTGTGTCTGT	3180
	CTGTGTGTAT	GCATACATAT	GTGTGTATAT	ATGTTTTTGT	CAGGTGTGTA	AATTGCAAAA	3240
	TTGTTTCTTT	TATATATGTA	TGTATATATA	TATATGAAAA	TATATATATA	TATGAAAAAT	3300
	AAAGCTTAAT	TGTCCAGAAA	AATCATACAT	TGCTTTTTTA	TTCTACATGG	GTACCACAGG	3360
	AACCTGGGGG	CCTGTGAATC	TACAACCAAA	AGGCACACAA	AACCGTTTCC	AGTTGGCAGC	3420
55	AGAGATCAGG	GGTTACCTCT	GCTTCTGAGC	AAATGGCTCA	AGCTTACCA	GAGCAGACAG	3480
	CTACCTTACT	TTTCAGCAGC	AAAACGTCCC	GTATGACGCA	GCACGAAGGG	CCTGGCAGGC	3540
	TGTTAGCAGG	AGCTATGTCC	CTTCTATCG	TTTCCGTCCA	CTT		

60 Seq ID NO: 120 Protein sequence:
Protein Accession #: NP_006491.1

	1	11	21	31	41	51	
65	MGLPRLVCAF	LLAACCCCP	VAGVPGEAEQ	PAPELVEVEV	GSTALLKCGL	SQSQGNLSHV	60
	DWFSVHKEKR	TLIFRVRRQG	QSEPEGEYEQ	RLSLQDRGAT	LALTQVTPQD	ERIFLCQGRK	120
	FRSQEYRIQL	RVYKAPSEPN	IQVNPFGIPV	NSKEPEEVAT	CVGRNGYPIP	QVIWYKNGRP	180
	LKEEKNRVIH	QSSQTVSSSG	LYTLQSLILKA	QLVKEDKDAQ	FYCELVNRLP	SGNHMKESRE	240
	VTVPVFYFTE	KVWLEVPVG	MLKEGDRVEI	RCLADGNPPP	HFSISKQNPB	TREAESETTN	300
70	DNGVLVLEPA	RKEHSGRYEC	QAWNLDTMIS	LLSEPEQLLV	NYVSDVRVSP	AAPERQEGSS	360
	LTLTCEABSS	QDLFPQWLRE	ETDQVLERGP	VLQLHDLKRE	AGGGYRCVAS	VPSIPGLNRT	420
	QLVRLAIFGP	PWMAFKERKV	WVKENMVLNL	SCEASGHFRP	TISWNVNGTA	SEQDQDPQV	480
	LSTLNVLVTP	ELLTGVVECT	ASNLDGKNTS	ILFLELVNLT	TLTPDSNTTT	GLSTSTASPH	540
75	TRANSTSTER	KLPEPESRGV	VIVAVIVCIL	VLAVLGAVLY	FLYKRGKLPK	RRSGKQBITL	600
	PFSRKTELTV	EVKSDKLPEE	MGLLQSSSGD	KRAPGDQGEK	YIDLRH		

80 Seq ID NO: 121 DNA sequence
Nucleic Acid Accession #: NM_018306
Coding sequence: 60-671

	1	11	21	31	41	51	
85	ATAGTCTACA	CAGAGCTCCC	CTTGTGCCCC	AGACAAGCTG	AAGGACCACA	GGAAAAAGCCA	60
	TGGAGACTTC	AGCATCTCC	TCCAGCCTC	AGGACAACAG	TCAAGTCCAC	AGAGAAACAG	120
	AAGATGTAGA	CTATGGAGAG	ACAGATTTC	ACAAGCAAGA	CGGGAAGGCT	GGACTCTTTT	180
	CCCAAGAAC	ATATGAGAGA	AACAAGTCTT	CTTCTCTCTC	CTTCTCTTCC	TCCTCATCCT	240
	CCTCATCTTC	TTTATCTCTC	TCCTCTCTCAG	GTCCTGGGCA	TGGGAGCCT	GACGTTTTGA	300

5
10
15
20

```
AGGATGAGCT TCAACTCTAT GGAGATGCTC CTGGAGAGGT GGTACCCCTCT GGGGAATCAG 360
GACTCCGAGG GAGAGGCTCT GACCCAGCAA GTGGAGAAGT GGAGGCTCT CAGTTAAGAA 420
GACTGAATAT AAAGAAGAT GATGAGTTT TCCATTTCGT CCTCTGTGC TTTGCCATCG 480
GGGCTTGCT GGTGTGTTAT CACTATTACG CAGACTGGTT CATGTCTCTT GGGGTGGGCC 540
TGCTCACCTT CGCCTCCCTG GAAACCGTTG GCATCTACTT CGGACTAGTG TACCGTATCC 600
ACAGCGTCTT CCAAGGCTTC ATCCCTCTCT TCCAGAAGTT TAGGCTGACA GGGTTCAGGA 660
AGACTGACTG AGGCCACTTC CAGGTGGGCA GCAGAGGCAG GCCCAGTGT GACCACCACT 720
GGGACCCCTG AGCCACACAAG GGCAGAGCAG CATTCTGAGA GACGCACAGG AGACCAAGCC 780
AGACCAATAA ACAGAACACT TTTCTTCCA TGTGGTCTGA ATGTTGGCAC CAGCCCGGGC 840
AGGGCATCTT CATTGGGCA GTACTGCTGT GCAACCCAGC TGCAAGGATG GAAGGCAGAG 900
GGTGGGTGTG GGGCCTGAGG CTTACAGTA CCTGGACCAG CAGGAAGATT CTGGAGGTC 960
ACTGCTCTCA GAGGACAGCA AGGGACCCCTG AGCTCTGCAA GCTGTGATCT GTCTGGGTTT 1020
ATGTTTTC TCAAATCCCA GGCATCTGCG ATGCGCTCTC AGGTGCTACC GAGCCATCCT 1080
GGGAGAGATG GATGTTCCAC TGCTTTGAGG CAGGGAGCCA TCGGGCTGGG GCCCTTGGT 1140
GAACCTGATG CAGGTAAGAT GCTGAGGACT AAAACCAATT TTTTGCACC CAAAAAATAA 1200
GGCAGGAAAA TGATCATCAG AAACATAATG GCAGCCAGGC ATGGGGGCTC ACGACTGTAA 1260
TCTCGCACT TTGGGAGGCT CAGGCTAAGG GTGCTTGAA GCTGAGAGTT CAAGACCAAC 1320
CTGGGCAACA TAGTGAGACC CCCATCTCTA CAATTTTTTT TTAATGACCA AATGTGGGG 1380
TACATACCTG TACATACCTG CGTTCCAGC TACTCAAGAG CGTGAGGCAG GAGGACTGCT 1440
TGAGCCAGG AGTTCAGGC TGCACTGAGG TAGCATCAAG CCACTGCATC CCAGCCTGGG 1500
CGACAGAGCA AGATCGTTTC TCTAAATTT
```

25 Seq ID NO: 122 Protein sequence:
Protein Accession #: NP_060776

30

```
1 11 21 31 41 51
METSASSSQP QDNSQVHRET EDVDYGETDF HKQDQKAGLF SQEQYERNKS SSSSFSSSSS 60
SSSSSSSSSS GPGHGEPDVL KDEQLQYQDA PGEVVPSEGS GLRRGSDPA SGEVEASQLR 120
RLNKKDDEF FHFVLLCPAI GALLVCYHY ADWFMISLGVG LLTFASLETV GIYFLGVYRI 180
HSVLQGFPL PQKFRLTGFR KTD
```

35 Seq ID NO: 123 DNA sequence
Nucleic Acid Accession #: BC022542
Coding sequence: 243..896

40
45
50
55
60
65
70

```
1 11 21 31 41 51
ACTTGGTCCC AGCCGATAAA TCTGGGGCAG CGCGCGGTAG GAGCTGCGGG CGGCCAGGCC 60
CCTTCTCTGG TCCGCACTCG GCCCGCGCGG CCCCTCTCGG GCGTCCGGCT TCCGGCGTCC 120
TGGCGGCTCG GTTGGCGGCG GTTGGCGGCG CGCCTCTGGT GCTCCTCGGG CGCGGCGACG 180
GGCTCAGCGG CGGGCCCGCC AGGCGCTTCA CGCGCGCGCG CTCTGACGCC GGCATAAGGG 240
CAGTGTGTTT TGAAATTATT TTGAGGCAAG AAGTTTGTAA AGATGGTTTC CACAGAGACC 300
TTTTAATCAA AGTGAAGTTT GGGGAAAGCA TTGAGGACTT GCACAGTGC CGTCTCTTAA 360
TTAAACAGGA CATTCTCTCA GGAATTATAT TGATTCOGTA TGAGTTGGCT TCATTACGAG 420
AGAGAAACAT AACAGAGGCA GTGATGGTTT CAGAAAATTT TGATATAGAG GCCCTAACT 480
ATTGTCCAA GGAGCTGAA GTTCTCATTT ATGCCAGACG AGATTACAG TGCAATTGACT 540
GTTTCAAGC GTTTTGGCT GTGCACTGCC GCTATCATCG CGCGCACAGT GAAGATGGAG 600
AAGCCTCGAT TTGGTCAAT AACCCAGATT TGTGATGTT TTGTGACCA GAGTCCCGA 660
TTTTGAAATG CTGGGCTCAC TCAGAAAGTG CAGCCCTCTG TGCTTTGGAT AATGAGGATA 720
TATGCCAATG GAACAGTATG AAGTATAAAT CAGTATATAA GAATGTGATT CTACAAGTTC 780
CAGTGGGACT GACTGTACAT ACCTCTCTAG TATGTTCTGT GACTCTGCTC ATTACAATCC 840
TGCTCTTAC ATTGATCTCT GTAGCAGTTT TCAAAATAGG CCATTTTTC CTATAAGTTT 900
TATGTAGTTA AATGCTTCT AGAAACCTAA ATAAGATCTA TTAATTTCTG ACGAGAGGTG 960
TTCTCTTAGA ATTAATTACT TTATCTTTT GTCTTCATT GTGGCCAAA TTAATTTTAC 1020
TAGAGGAAT TTGGGATCAT TCTCAGCTAA TTCCAAAATG TAGTGTCTTA TTGCATGGAT 1080
CCTGGTAAAT CCTCAGCAT CAGATGCCAT AAGGGGAAAC TTAATTCTGC TAAATTAATG 1140
TTTATTTGT GAGAAGTGAC TTTATCTTCA TTGGGGTAG AAAAATTATT TCTTTATGTA 1200
GTAGAGACAA ATTATTCTCA TTTTGCAAGT ACTTTCAATT TAAGCTACAA ATTGAGAAAA 1260
CGGTATATAA TAAGAAATAA ATAGGCCAGG CACAGTGGCT CACACCTGTA ATCCAGCAC 1320
TTTGGGAGGC CGAGGTGGGC GGATCACCAG AGGTCAAGAG TTTGAGACCA GCTTGGTGAA 1380
ACCTGTCTCT TACTAAAAAT ACAAAAGTTA GCTGGGGCTG GTGGTGGGCA TCTGTAGTCC 1440
CAGCTAATTG GAAGGGTGAG CGCGGAGGAT CGCTTGAACC TGGGAGGCGG AGGTTCCAGA 1500
GAGCCAAGAT CGCACCCTG CACTACAGCC TGGGCGACAG AACGAGACCC TGCTCCAAA 1560
GGAAAAACAA AAAAGAGAA TAAATAATT TGATGAAAA TCATGTTTAT TTAATAGTA 1620
ATGTCAAGAG ACTATTAAAG ATGTGCCAGA GTTCAATGA AAATCATTA AGTAGGACAG 1680
CTAAGAAATT AATATTAAATA TAAAAATTAT TGATAATCTT AAATTATGTA TTATCTCTA 1740
ACGCACTCCA TTCTCTTTT ACATTTTATC ATGTTTCTTT TGAATATATG AATTGGCAAA 1800
GGACTTGATG AAACCTGAGT CTAAGATTG GTACAGAGTA TGTCAGGAAG ACAACTCAGA 1860
TTGCCATTTT AAATAAAGT GTACATGAAC AAAAAAATAA AAAAAA
```

75 Seq ID NO: 124 Protein sequence:
Protein Accession #: AAH22542

80

```
1 11 21 31 41 51
MCSEIILRQE VLKDGPHRDL LIKVKFGESI EDLHTRLLI KQDIPAGLYV DPYELASLRE 60
RNITEAVMVS ENFDIEAPNY LSKSEVLIY ARDSQCIDC FQAFLPVHCR YHRPHSEDE 120
ASIVVNNPDL LMFCDQAGSR RMIRFRFDSF DKTIEFFILK CWAHSEVAAP CALENEDICQ 180
WNMKYKSVY KNVILQVPVG LTVHTSLVCS VTLLITILCS KRKKK
```

85 Seq ID NO: 125 DNA sequence
Nucleic Acid Accession #: NM_004994.1
Coding sequence: 20..2143

1 11 21 31 41 51
AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCTG GTCTCTGGTG TCCTGGTGCT 60
GGGCTGCTGC TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTGTGCTCT TCCTGGGAGA 120
5 CCTGAGAAC AATCTCACCG ACAGGCGAGT GGCAGAGGAA TACCTGTACC GCTATGGTTA 180
CACTCGGGTG GCAGAGATGC TGGAGAGTGC GAAATCTCTG GGGCTGCGC TGCTGCTTCT 240
CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300
GGGAACCCCA CGGTGGGGGG TCCAGACCTT GGCAGATTC CAACCTTTTG AGGGCGACCT 360
10 CRAAGTGGCAC CACCAACAAC TCACCTATTG GATCCAAAG TACTCGAAG ACTTGCOCG 420
GGCGGTGATT GACGAGCCTT TTGCCGCGC CTTCGCACTG TGGAGCGCGG TGACGCGCCT 480
CACCTTCACT CGGTGTATCA GCGGGAGCG AGACATGCTC ATCCAGTTTG GTGTGCGGA 540
GCACGGAGAG GGGTATCCCT TOGACGGGAA GGACGGGCTC CTGGCACAGC CTTTTCCTCC 600
TGGCCCCGGC ATTCAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCTGGGCAA 660
15 GGGCGTGGC GTTCCAACCT GGTTTGGAAG CGCAGATGGC GCGGCTGCGC ACTTCCCTTT 720
CATCTTCGAG GGCCTCTCT ACTCTGCTG CACCAACGAC GGTGCTCGC ACGGCTTGCC 780
CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGCTTCT GCCCAGCGA 840
GAGACTCTAC ACCCGGGAGC GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT 900
CCAAGGCGAA TCTACTCCG CCGTCAACAC GGAOGTTCGC TCCGACGGCT ACCGCTGGTG 960
20 CGCCACCAAC GCGCACTAG ACGGGGACAA GCTCTTCGGC TCTGCGCGA CCGGAGCTGA 1020
CTGACGCGTG ATGGGGGCGA ACTCGGGGGG GGAGCTGTGC GTCTTCCCTC TCACTTTCCT 1080
GGGTAAAGAG TACTCGACCT GTACCGGCGA GGGCGCGCGG GATGGGCGCC TCTGGTGGC 1140
TACCACCTCG AACTTTTGAC GCGACAAGAA GTGGGCTTC TGCCCGGACC AAGGATACAG 1200
TTTGTTCCTC GTGGCGGGCG ATGAGTTGCG CCAOGCGCTG GGTATGATC ATTCTCTCAGT 1260
25 GCGGAGGCG GTTCACTAC CTATGTACCG CTTCACTGAG GGGCCCCCTC TGCAATAGGA 1320
CGACGTGAAT GGCATCGGC ACCTCTATGG TCTCGCCCTC GAACCTGAGC CACGGCTCC 1380
AACCACCAAC ACACCGCAGC CACCGGCTCC CCGACGGTTC TGCCCCACCG GACCCCCCAC 1440
TGTCCACCCG TCAGAGCGCC CCACAGCTGG CCCCACAGT CCCCCCTCAG CTGGCCCCAC 1500
AGGTCCCCCG ACTGTGCGCC CTTCTACGGC CACTACTGTG CTTTGTAGTC CGGTGGACGA 1560
30 TGCTGCGAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGACACGAG TGTATTGTGT 1620
CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGGG AGCGCGCGCG AGGGCCCCCT 1680
CCTTATCGCC GACAAGTGGC CCGCGCTGCC CCGCAGCTG GACTCGGTCT TTGAGGAGCC 1740
GCTCTCCAG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGGTC 1800
GGTGTCTGGC CCGAGCGCTC TGGACAAGCT GGGCCTGGGA GCGCAAGTGG CCGAGGTGAC 1860
35 CGGGGCCCTC CGGAGTGGCA GGGGGAAGAT GCTGCTGTT AGCGGGCGCG GCCTCTGAG 1920
GTTTCGACGTG AAGCGCGAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGATGTT 1980
CCCCGGGGTG CCTTTGGACA CGCACGAGCT CTTCAGTAC CGAGAGAAAG CCTATTTCGT 2040
CCAGACCGC TCTACTGGC CGGTGAGTTC CCGGAGTGAG TTGAACAGG TGGACCAAGT 2100
GGGCTACGTG ACCTATGACA TCTGCAAGT CCGTGAAGAC TAGGGCTCCC GTCTCTCTTT 2160
40 CAGTGGCAT GTAATCCCG ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCGGATA 2220
CAAACCTGTA TTCTGTTCTG GAGGAAAGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC 2280
TCACCTTGT TTTTGTGTG AGTGTCTTA ATAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 126 Protein sequence:
Protein Accession #: NP_004985.1

1 11 21 31 41 51
MSLWQPLVLV LLVLGCCFAA PRQRQSTLVL FPGDLRLNLT DRQLAEYLY RYGYTRVAM 60
RGESKLGPA LLLKQLSL PETGELDSAT LKAMRTPRCG VPDLRFPOTF EGDWKWHHN 120
50 ITYWIQYSE DLPRVIDDA FARAPALWSA VTPLTFTRVY SRDADIVIQF GVAEHGDGYP 180
FDKGDLGLAH APPFGPIQG DAHFDDELW SLGKGVVPT RFGNADGAAC HPPFIFEGRS 240
YSACTDGRS DGLPWCSTTA NYDTDDRFGF CPSERLYTRD GNADGKPCQF PFIFQGSYS 300
ACTTDRSGD YRWCAATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTFLGKEYST 360
55 CTSEGRGDR LWCATTSNFD SKKMGFCPD QGYSLEPLVAA HEPHALGLD HSSVPEALMY 420
PMYRFTGPP PHKDDVNGIR HLYGPRPEPE PRPFTTTTPQ PTAPPTVCPT GPPTVHPSER 480
PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIABIQNG LYLFLKDGKYW 540
RFSERGRSR QGFLLADKW PALPRKLDV FEEPLSKLF PFSGRQVWVY TGAIVLGRPR 600
60 LDKLGLGADV AQVTGALRSR RGKMLLFSR RLNRFDVKAQ MVDPRSASEV DRMFPGVPLD 660
THDVQYREK AYPCQDRFYW RVSSRSELAQ VDQGVVYTYD ILQCPED

Seq ID NO: 127 DNA sequence
Nucleic Acid Accession #: NM_004181
Coding sequence: 32-670

1 11 21 31 41 51
GCAGAAATAG CTAAGGAGA TCAACCCGA GATGCTGAAC AAAGTGTCTG CCCGGCTGGG 60
GGTCGCGGCG CAGTGGCGCT TCGTGGACGT GCTGGGCTG GAAGAGGAGT CTCTGGGCTC 120
70 GGTGCCAGCG CTGCGCTGCG CGCTGCTGCT GCTGTTTCCC CTCACGGCCC AGCATGAGAA 180
CTTCAGGAAA AAGCAGATTG AAGAGCTGAA GGGACAAGAA GTTAGTCCTA AAGTGTACTT 240
CATGAAGCAG ACCATTGGGA ATTCTGTGG CACAATCGGA CTTATTCAAG CAGTGCCCAA 300
TAATCAAGAC AAACCTGGGAT TTGAGGATGG ATCAGTTCTG AAACAGTTTC TTTCTGAAAC 360
AGAGAAAATG TCCCTGAAG ACAGAGCAA ATGCTTTGAA AAGAATGAGG CCATACAGGC 420
75 AGCCCATGAT GCGTGGCAC AGGAAGGCCA ATGTCGGGTA GATGACAAGG TGAATTTCCA 480
TTTATTCTG TTTAACAACG TGGATGGCCA CCTCTATGAA CTTGATGGAC GAATGCCTTT 540
TCCGGTGAAC CATGGCGCCA GTTCAGAGGA CACCTGCTG AAGGACGCTG CCAAGGTGTG 600
CAGAGAAATC ACGAGCGTG AGCAAGGAGA AGTCCGCTTC TCTGCGGTGG CTCTCTGCAA 660
GGCAGCCTAA TGCTCTGTGG GAGGGACTTT GCTGATTTC CCTCTTCCC TCAACATGAA 720
AATATATACC CCGATGCGAG TCTAAAATGC TTCAGTACTT GTGAAACACA GCTGTCTCTC 780
80 TGTCTGCGAG ACACCGCTTC CCTCAGCCA CACCCAGGCA CTTAAGCACA AGCAGAGTGC 840
ACAGCTGTC ACTGGGCCAT TGTGGTGTGA GCTTCAGATG GTGAAGCATT CTCCCCAGTG 900
TATGCTTGT ATCGATATC TAACGCTTA AATGGCTACT TTGCTTCTG TCTGTAAAGT 960
AAGACCTTGG ATGTGGTTAT GTTGCTCTTA AGAATAAATT TTGCTGATAG TAGC

Seq ID NO: 128 Protein sequence:
Protein Accession #: NP_004172

1 11 21 31 41 51
 5 MLNKVLSRIG VAGQWRFDV LGLEESLGS VPAPACALLL LFPLTAQHEN FRKKQIEELK 60
 GQEVSPKVYP MKQTIGNSG TIGLIHAVAN NQDKLGFEDG SVLKQFLSET EKMSPEORAK 120
 CFENEAIGA AHDAVAQEGQ CRVDDKVNPH FILFNNVDGH LYELDGRMPF PVNHGASSED 180
 TLLKDAAKVC REPTEREQGB VRFSVALCK AA

10 Seq ID NO: 129 DNA sequence
 Nucleic Acid Accession #: NM_000213
 Coding sequence: 127-5385

1 11 21 31 41 51
 15 CGCCCGCGCG CTGCAGCCCC ATCTCCTAGC GGCAGCCGAG GCGCGGAGGG AGCGAGTCGG 60
 CCCCAGGGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GGCTGGCCGG GAGAGGGAGG 120
 AAGAGGATGG CAGGGCCACG CCCAGGCCCA TGGGCCAGGC TGCTCTGGGC AGCCTTGATC 180
 AGCGTCAGCC TCTCTGGGAC CTTGGCAAAAC CGCTGCAGA AGGCCCCAGT GAAGAGCTGC 240
 20 ACGGAGTGTG TCCGTGTGGA TAAGGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC 300
 CGGCGCTGCA ACACCCAGCG GGAGCTGTCT GCGCGGGGCT GCGAGCGGGA GAGCATCGTG 360
 GTCATGGAGA GCAGCTTCCA AATCACAGAG GAGACCCAGA TTGACACCCG CCTGCGGGCG 420
 AGCCAGATGT CCCCCCAAGG CCTGCGGGTC CGTCTGCGGC CCGGTGAGGA GCGGCATTTT 480
 GAGCTGGAGG TGTTTGAGCC ACTGGAGAGC CCGGTGGACC TGTACATCCT CATGGACTTC 540
 25 TCCAACTCCA TCTCCGATGA TCTGGACAAC CTCAGAAGA TGGGGCAGAA CCTGGCTCGG 600
 GTCTGAGCC AGCTCACCAG CGACTACACT ATTGGATTGT GCAAGTTTGT GGACAAAGTC 660
 AGGTGCCCGC AGACGACAT GAGGCTGAG AAGCTGAAGG AGCCTGGCC CAACAGTGAC 720
 CCCCCCTTCT CTTTCAAGAA CGTCATCAGC CTGACAGAAG ATGTGGATGA GTTCCGGAAT 780
 AAACCTCAGG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CTGAGGGCGG CTTGGATGCC 840
 30 ATCTGCAGA CAGCTGTGTG CAGGAGGAC ATTGGCTGGC GCGCGGACAG CACCCACCTG 900
 CTGGTCTTCT CCAACGAGTC AGCCTCCAC TATGAGGCTG ATGGCGCAA CGTGCTGGCT 960
 GGCAATCAGA GCGCAACCA GTAAACGGTC CACCTGGACA CCAAGGGCAC CTACACCCAG 1020
 TACAGGACAC AGGACTACCC GTCCGTGCCC ACCCTGGTGC GCCTGCTCGC CAAGCACAAC 1080
 ATCATCCCCA TCTTTGCTGT CACCACTAC TCCTATAACT ACTACGAGAA GCTTCAACAC 1140
 35 TATTTCTCTG TCTCCTCACT GGGGGTGTG CAGGAGGACT CGTCCAACAT CTGGAGCTG 1200
 CTGGAGGAGG CTTTCAATCG GATCCGCTCC AACCTGGACA TCGGGCCCT AGACAGCCCC 1260
 CGAGGCTCTT GGACAGAGGT CACCTCCAAG ATGTTCCAGA AGACGAGGAC TGGTCTCTTT 1320
 CACATCGGCG GGGGGGAAGT GGGTATATAC CAGGTGCAGC TGGGGCCCT TGAGCAGGTG 1380
 GATGGGACGC ACGTGTGCCA GCTGCCGAG GACCAGAAAG GCAACATCCA TCTGAACCT 1440
 40 TCTTCTCTCG ACGGCTCAA GATGGACGCG GGCATCATCT GTGATGTGTG CACCTGCGAG 1500
 CTGCAAAAG AGTGTGGGTC AGCTCGCTGC AGCTTCAAAG GAGACTTCGT GTGCGGACAG 1560
 TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT 1620
 GACATTCAGC CTTGCCCTGC GAGAGGGGAG GACAAAGCGT GCTCGGGCCG TGGGAGGTG 1680
 CAGTGGGGCG ACTGTGTGTG CTACGGGAG GCGCGCTACG AGGGTCAGTT CTGCGAGTAT 1740
 45 GACAACTCG AGTGTCCCCG CACTTCCGGG TTCTCTGCA ATGACCGAGG ACGCTGCTCC 1800
 ATGGGCCAGT GTGTGTGTGA GCCTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCTCAGC 1860
 AATGCCACTT GCATCGACAG CAATGGGGGC ATCTGTAATG GACGTGGCCA CTGTGAGTGT 1920
 GCGCGCTGCG ACTGCCACCA GCAGTGGCTC TACACGGACA CCATCTGCGA GATCAACTAC 1980
 TCGCGATGCC ACCCGGGCCT CTGCGAGGAC CTACGCTCCT GGTGCAAGT CCAGGCGTGG 2040
 50 GGCACGGGCG AGRAGAGGG GCGCACTGT GAGGAATGCA ACTTCAAGGT CAAGATGGTG 2100
 GACGAGCTTA AGAGAGCCGA GGAGGTGGTG GTGCGCTGCT CCTTCCGGGA CGAGGATGAC 2160
 GACTGCACTT ACAGTACAC CATGGAAGGT GACGGGCCCC CTGGGCCCAA CAGCACTGTC 2220
 CTGGTGACA AGAAGAACCA CTGCCCTCGG GCTCCTTCT GGTGGCTCAT CCCCCTGCTC 2280
 CTCTCTCTCC TGGCGCTCCT GGCCCTGCTA CTGTGCTAT GCTGGAAGTA CTGTGCTGTC 2340
 55 TGCAGGCGCT GCTTGGCACT TCTCCGTCG TGCAACGAG GTCACATGTT GGGCTTTAAG 2400
 GAAGCACTT ACATGCTGCG GGAGAACCTG ATGGCTCTG ACCACTTGA CAGCCCCATG 2460
 CTGCGCAGCG GGAACCTCAA GGGCGGTGAC GTGGTCCGCT GGAAGGTAC CAACAACATG 2520
 CAGCGGCTCG GCTTTGCCAC TCATGCGGCC AGCATCAACC CCACAGAGCT GGTGCCCTAC 2580
 GGGCTGTCTT TGGCCTTGCG CCGCTTTGCG ACCGAGAAC TGCTGAAGCC TGCACTCGG 2640
 60 GAGTGGCGCC AGTGTGCGCA GGAGGTGGAG GAGAACCTGA ACGAGTCTA CAGGCAGATC 2700
 TCGGTGTGAC ACAAGCTCCA GCAGACCAAG TTCCGGCAGC AGCCCAATGC CGGAAAAAG 2760
 CAGAGCCACA CCATTGTGGA CACAGTGTCT ATGGGCCCC GCTCGGCCAA GCGCGCCCTG 2820
 CTGAAGCTTA CAGAGAAGCA GGTGGAACAG AGGGCCTTCC ACGACCTCAA GGTGGCCCCC 2880
 GGTCTACTACA CCCTCACTCG AGACAGGAC GCGCGGGGCA TGGTGAAGTT CCAGGAGGGC 2940
 65 TGGAGCTCG TGGAGCTAGC GGTGCCCTTC TTTATCCGGC CTGAGGATGA CAGCAGAAAG 3000
 CAGCTGCTGG TGGAGGCCAT CGAGTGCCCC GCAGGCACTG CCACCTCGG CCGCGGCTG 3060
 GTAAACATCA CCATCATCAA GGAGCAAGCC AGAGACGTGG TGTCTTTGA GCAGCCTGAG 3120
 TTCTCGGTCA GCGCGGGGA CCAGGTGGCC CGCATCCCTG TCATCCGGCG GTCTCTGGAC 3180
 GGGCGGAAGT CCGAGGTCTC CTACCGCACA CAGGATGGCA CCGCGCAGGG CAACCGGGAC 3240
 70 TACATCCCCG TGGAGGGTGA GCTGCTGTTT CAGCCTGGGG AGGCCTGGA AGAGCTGCAG 3300
 GTGAAGCTCC TGGAGCTGCA AGAAGTTGAC TCCCTCTGCG GGGGCGGCCA GGTCCGCGGT 3360
 TTCCAGTCC AGCTCAGCAA CCCTAAGTTT GGGGCCACCC TGGGCCAGCC CCATCCACCC 3420
 ACCATCATCA TCAGGGACCC AGATGAACTG GACCGAGCT TCACGAGTCA GATGTTGTCA 3480
 TCACAGCCAC CCGCTCAGCG CGACCTGGGC GCGCGGAGA ACCCAATGC TAAGSCCGCT 3540
 75 GGTCCAGGA AGATCCATT CAACCTGGCTG CCCCCTTCTG GCAAGCCAT GGGGTACAGG 3600
 GTAAAGTACT GGATTGAGG TGACTCGAA TCCGAAGCCC ACCTGCTCGA CAGCAAGGTG 3660
 CCCTCAGTGG AGCTCACCAG CCTGTACCCG TATTGCGACT ATGAGATGAA GGTGTGCGCC 3720
 TACGGGCTCT AGGCGAGGG ACCCTACAGC TCCCTGGTGT CCGCGCGCAC CCACAGGAA 3780
 GTGCCAGCG AGCCAGGGGG TCTGGCCTTC AATGTCTCT CCTCCAGGT GACCCAGCTG 3840
 80 AGCTGGGCTG AGCGCGCTGA GACCAACGGT GAGATCACAG CTTACGAGGT CTGCTATGGC 3900
 CTGGTCAAG ATGACAAAGC ACCTATTGGG CCGATGAAGA AAGTGTGTT TGACAACTCT 3960
 AAGAACCGGA TGCTGCTTAT TGAGAACCTT CCGGAGTCCC AGCCCTACCG CTACAGGCTG 4020
 AAGCGCGCA ACGGGGCGCG CTGGGGCCTT GAGCGGGAGG CCATCATCAA CCGTGGCCAC 4080
 CAGCCCAAGA GGGCCATGTC CATCCCATC ATCCCTGACA TCCCTATCGT GGACGCCAG 4140
 85 AGCGGGGAGG ACTACGACAG CTTCCTTATG TACAGCATG ACGTCTTACG CTCTCCATCG 4200
 GGCAGCCAGA GGCAGGAGG CTCCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGGAC 4260
 TTTGCTCTCC CGGCGAGCAC CAACCTCCCTG CACAGGATGA CCAAGACAG TGCTGCTGCC 4320
 TATGGCACCC ACCTGAGCCC ACAGTGCCCC CACGCGGTGC TAAGCACATC CTCCACCTCT 4380

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

ACACGGGACT ACAACTCACT GACCCGCTCA GAACACTCAC ACTCGACCAC ACTGCCGAGG 4440
GACTACTCCA CCTCTACCTC CGTCTCCTCC CACGACTCTC GCCTGACTGC TGGTGTGCC 4500
GACACGCCCA CCGCCTCGGT GTTCTCTGCC CTGGGGCCCA CATCTCTCAG AGTGAGCTGG 4560
CAGGAGCCGC GGTGCGAGCG GCGCTGCAG GGCCTACAGT TGGAGTACCA GCTGCTGAAC 4620
GGCGGTGAGC TGCATCGGCT CAACATCCCC AACCTGCCCC AGACCTCGGT GGTGGTGGAA 4680
GACCTCCTGC CCAACCACTC CTACGTGTTT CCGGTGCGGG CCCAGAGCCA GGAAGGCTGG 4740
GGCCGAGAGC GTGAGGSGTG CATCACCATT GAATCCCAGG TGCACCCGCA GAGCCCACTG 4800
TGTCCCTCGC CAGGCTCCGC CTTCACCTTG AGCACTCCCA GTGCCCCAGG CCCGCTGGTG 4860
TTCACTGCCC TGAGCCCAAG CTGCTGTCAG CTGAGCTGGG AGCGGCCACG GAGGCCCAAT 4920
GGGATATCG TCGGCTACCT GGTGACCTGT GAGATGSCCC AAGGAGGAGG GCCAGCCACC 4980
GCATTCCGGG TGGATGGAGA CAGCCCCGAG AGCCGGCTGA CCGTGCCGGG CCTCAGCGAG 5040
AACGTGCCCT ACAAGTTCAA GGTGCAGGCC AGGACCACTG AGGCTTTCGG GCCAGAGCCG 5100
GAGGGCATCA TCACCATAGA GTCCCAAGAT GGAGGACCTT TCCGCGAGCT GGGCAGCCGT 5160
GCGGGCTCTC TCCAGCACCC GCTGCAAAAG CAGTACAGCA GCATCACCA CACCCACACC 5220
AGGCCACCG AGCCCTCTCT AGTGGATGGG CCGACCTCGG GGGCCACGCA CCTGGAGGCA 5280
GGCGGCTCCC TCACCCGAGA TGTGACCCAG GAGTTTGTGA GCGGACACT GACCACGAGC 5340
GGAACCCCTA GCACCCACAT GGACCAACAG TTCTTCCAAA CTTGACCGCA CCTGCCCA 5400
CCCCGCCAT GTCCCACTAG GCGTCTCTCC GACTCTCTTC CCGGAGCCTC CTCAGCTACT 5460
CCATCTCTGC ACCCTGGGG GCGCCAGCCA CCGCATGCA CAGAGCAGGG GCTAGGTGTC 5520
TCCTGGGAGG CATGAAGGGG GCAAGGTCCG TCCTCTGTGG GCCCAAACCT ATTTGTAACC 5580
AAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTCTG CACTTAATAA ATGTTTGTG 5640
ACTG

Seq ID NO: 130 Protein sequence:
Protein Accession #: NP_000204

1 11 21 31 41 51
30 MAGPRPSPWA RLLAALISV SLGTLANRC KKAPVKSCTE CVRVKDCAY CTDEMPFRDR 60
CNTQAELLAA GQCRESVVM ESSFQITEET QIDTTLRRSQ MSPQGLRVRL RPGEERHFEL 120
EVPEPLESPV DLVILMDPSN SMSDDLNLK KMGQNLARVL SOLTSDYITG FGKFPVKVSV 180
PQTDMPREKL KEPWPNSDPP FSPQNVISLT EDVDEPRNKL QGERISGNLD APEGGPDAIL 240
QTAVCTRDIG WRDPSLHLIV FSTESAFHYE ADGANVLAGI MSRNDERCHL DTTGTYTOYR 300
35 TQDYPSVPTL VRLAKENII PIFAVTNYSY SYYEKLHYE FVSSLGLVLE DSSNIVELLE 360
EAFNRIRSNL DIRALDSPRG LRTEVTSKMF QKTRTGSFHI RRGVSGIYQV QLRALHVDG 420
THVQQLPEQ KGNHILKPSF SDGLKMDAGI ICDVCTCELQ KEVRSARCSF NGDFVCGQCV 480
CSBWSGGTC NCSTGSLSDI QPCLRBEGDK PCSGRGECQC GHVCYGBGR YEGQFCEYDN 540
40 FQCPRTSGFL CNDRGRCMSG QCVCEPGWTG PSCDCPLSNA TCIDSNGGIC NGRGHCECGR 600
CHCQQSLYT DTICEINYSY IHPGLCEDLR SCVQCQAWGT GEKKGRTECE CNFKVKMVE 660
LKRAEVEVVR CSFRDEDDC TSYTMEGDG APGPNSTVLV HKKDCPPGS FWWLIPLLL 720
LLPLALLLL LCWKYCACCK ACLALLPCCN RGHMVGFKEH HYMLRENLM SAHLDTPLMR 780
SGNLKGRDVV RWKVTNMQR PGFATHAASI NPTELVPYGL SLRLARLCTE NLLKPDTR 840
45 AQLRQVEVEN LNEVYRQISG VHKLQTKPR QPNAGKKQD HTIVDTVMA PRSAKALLK 900
LTEKQVEQRA PHDLKAPGY YTLADQDAR GMVEFQEGVE LVDVRVPLFI RPEDDEKQL 960
LVEAIDVPAQ TATLGRRLVN ITIIEKQARD VVSFEQPEFS VSRGDQVARI PVIRRVLDGG 1020
RSQVSYRTDQ GTAQGNLDYI PVSGELLFQP GEANKELQVK LLELQEVDSL LRGRQVRRF 1080
VOLSNKFGA HLQGPSTTI IIRDFELDR SFTSOMLSQ PPHGDLGAP QNPNAKAAGS 1140
50 RKIHFWNPLP SGKPMGYRVK YWIQDSESE ABLDSKVPS VELTNLYPYC DYEMKVCAYG 1200
AQEGGPYSSL VSCRTHQEVV SEPGRALFNV VSSVTQLSW AEPATNGEI TAYEVCYGLV 1260
NDNDRPIGPM KKVLDNPNK RMLLIENLRE SQPYRYTVKA RAGAGWGPFR EAINLATQP 1320
KRPMIPIIP DIPIVDAQSG EDYDSFLMYS DDVLRSPSGS QRPVSVDDE HLNVGRMDFA 1380
55 PFGSTNSLHR MTTTAAAYG THLSPHVPHR VLSTSTLTR DYNLSLTRSEH SHSTTLPRDY 1440
STLTSVSSHD SRLTAGVPDT PTRLVFSALG PTLRVSQWE PRCEPLQGY SVEYQLNGG 1500
ELHRLNIPNP AQTSVVVEDL LPNHSYVFRV RAQSQEGNGR EREGVITIES QVHPQSPLCP 1560
LPGSAFTLST PSAPGLVFT ALSPDSLQLS WERPRRPNGD IVGYLVTCM AQGGGPATAP 1620
60 RVDGDSPEGR LTVPGLSNV PYKFKVQART TEGFQPEREG IITIESQDGG PFPQLSGRAG 1680
LFQHPQLQSEY SSITTTTSTA TEPFLVDGPT LGAQHELAGG SLTREVTQEF VSRTLTSGT 1740
LSTHMDQFF QT

Seq ID NO: 131 DNA sequence
Nucleic Acid Accession #: BC004372
Coding sequence: 132..2231

65 1 11 21 31 41 51
70 CCTCGTGGCG CGGACCCGAG CCTCTGCCAG GTTCGGTCCG CCATCCTCGT CCGTCTCTCC 60
GCCGCGCCCT GCCCGCGGCC CAGGGATCCT CCAGCTCCTT TCGCCCGGCG CCGCTCGTCC 120
CTCGGACAC CATGGACAAG TTTTGGTGGC ACGCAGCCTG GGGACTCTGC CTGCTGCCCG 180
75 TGAGCCTGGC GCAGATCGAT TTGAATATAA CCTGCCGCTT TGCAGGTGTA TTCCACGTGG 240
AGAAAAATGG TCGCTACAGC ATCTCTCGGA CGGAGGCGCG TGACCTCTGC AAGGCTTTCA 300
ATAGCACTTT GCCCACAATG GCCCAGATGG AGAAAGCTCT GAGCATCGGA TTTGAGACCT 360
GCAGGTATGG GTTCATAGAA GGGCATGTGG TGATTCCCCG GATCCACCCC AACTCCATCT 420
GTGCAGCAAA CAACACAGGG GTGTACATCC TCACATCCAA CACCTCCAG TATGACACAT 480
80 ATTGCTTCAA TGCTTCAGT CCACCTGAAG AAGATTGTAC ATCAGTCACA GACCTGCCCA 540
ATGCTTTTGA TGGACCAATT ACCATAACTA TTGTTAACCG TGATGGCACC CGCTATGTCC 600
AGAAAGGAGA ATACAGAAAG AATCCTGAAG ACATCTACCC CAGCAACCT ACTGATGATG 660
ACGTGAGCAG CGGCTCTCTC AGTGAAAGGA GCAGCAGTTC AGGAGGTTAC ATCTTTTACA 720
CTTTTCTCAT TGATACCCCC ATCCAGACAG AAGACAGTCC CTGGATCACC GACAGCAGAG 780
85 ACAGAAATCC TGCTACCACT ACCTCTTCAA ATACCATCTC AGCAGGCTGG GAGCCAAATG 840
AAGAAATGA AGATGAAAGA GACAGACACC TCAGTTTTC TGGATCAGGC ATTGATGATG 900
ATGAAGATT TATCTCCAGC ACCATTTCAA CCACACCAAG GGCTTTTGAC CACACAAAC 960
AGAACCAAGA CTGGACCCAG TGGAAACCAA GCCATTCAA TCCGGAAGTG CTACTTCAGA 1020
CAACCAAG GATGACTGAT GTAGACAGAA ATGGACCAAC TGCTTATGAA GGAAGTGA 1080
ACCCAGAAC ACCCTCTCCC CTCATTCAAC ATGAGCATCA TGAGGAAGAA GAGACCCAC 1140
ATTCTACAG CACATCCAG GCAACTCTA GTAGTACAAC GGAAGAAACA GCTACCCAGA 1200
AGGAACAGTG GTTTGGCAAC AGATGGCATG AGGGATATCG CCAACACCCC AGAGAAGACT 1260

5
10
15
20

```

CCCATTCGAC AACAGGGACA GCTGCAGCCT CAGCTCATAC CAGCCATCCA ATGCAAGGAA 1320
GGACAACACC AAGCCCAGAG GACAGTTCCT GGACTGATTT CTTCACCCCA ATCTCACACC 1380
CCATGGGAGC AGGTCATCAA GCAGGAAGAA GGATGGATAT GGACTCCAGT CATAGTACAA 1440
CGCTTCAGCC TACTGCAAAAT CCAAAACACAG GTTTGGTGGG AGATTGGGAC AGGACAGGAC 1500
CTCTTTCAAT GACAAAGCAG CAGAGTAATT CTCAGAGCTT CTCTACATCA CATGAAGGCT 1560
TGGAGAAGA TAAAGACCAT CCAACAACTT CTACTCTGAC ATCAAGCAAT AGGAATGATG 1620
TCACAGGTGG AAGAAGAGAC CCAATCAATT CTGAAGGCTC AACTACTTTA CTGGAAGGTT 1680
ATACCTCTCA TTACCCACAC ACGAAGGAAA GCAGGACCTT CATCCCAGTG ACCTCAGCTA 1740
AGACTGGGTC CTTTGGAGTT ACTGCAGTTA CTGTTGGAGA TTCCAACCTCT AATGTCAATC 1800
GTTCTTATC AGGAGACCAA GACACATTCC ACCCCAGTGG GGGGTCCCAT ACCACTCATG 1860
GATCTGAATC AGATGGACAC TCACATGGGA GTCAAGAAGG TGGAGCAAAC ACAACCTCTG 1920
GTCCTATAAG GACACCCCAA ATTCCAGAAAT GGCTGATCAT CTGCGCATCC CTCTTGGCCT 1980
TGGCTTTGAT TCTTGCAGTT TGCAATGCGA TCAACAGTCG AAGAAGGTGT GGGCAGAAGA 2040
AAAAGCTAGT GATCAACAGT GGCAATGGAG CTGTGGAGGA CAGAAAGCCA AGTGGACTCA 2100
ACGGAGAGGC CAGCAAGTCT CAGGAAATGG TGCATTGGT GAACAAGGAG TCGTCAGAAA 2160
CTCCAGACCA GTTTATGACA GCTGATGAGA CAAGGAACCT GCAGAATGTG GACATGAAGA 2220
TTGGGGTGTA ACACCTACAC CATTATCTTG GAAAGAAACA ACCGTTGGAA ACATAACCAT 2280
TACAGGGAGC TGGGACACTT AACAGATGCA ATGTGCTACT GATTGTTTCA TTGCGAATCT 2340
TTTTTAGCAT AAAATTTTCT ACTCTTAAAA AAAAAAAAAA AAAAAAAA

```

Seq ID NO: 132 Protein sequence:
Protein Accession #: AAH04372

25
30
35
40

```

1 11 21 31 41 51
MDKFWHAAM GLCLVPLSLA QIDLNITCRF AGVPHVEKNG RYISISRTEA DLCKAFNSTL 60
PTMAQMEKAL SIGFETCRYG FIEGHVVIPR IHPNSICAAH NTGVYILTSN TSQYDTCFNP 120
ASAPPEEDCT SVTDLENARD GPITITIVNR DGRYVQKGE YRTNPEDIYP SNPTDDDVSS 180
GSSSSSSSTS GGYIFYTFST VHIPDEDSF WITDSTDRIP ATSTSSNTIS AGWEPNEENE 240
DERDRHLSFS GSGIDDDDEF ISSTISTTPR AFDHTKQND WTQWNPESHN PEVLQTTTR 300
MTDVRNGTIT AYSGNWNPEA HPPLIHHEHH EBEETPHSTS TIQATPSSTT EETATQKEQW 360
FGNRWHEGYR QTPREDSSHST TGTAAASAHT SHPMQGRITP SPEDSSWTFD FNPISHPMGR 420
GHQAGRMDM DSSHSTTLQF TANFNTGLVE DLDRTGLPLM TTQSSNSQSF STSHEGLEED 480
KDHPTTSLT SNNRNDVTGG RRDPNHSEGS TTLLEGYTSR YPHTKESRTF IPVTSAGTGS 540
FGVTAVTVGD SNSNVNRSLS GDQDTFHPG GSHTTHGSES DGHSHGSEQB GANTTSGPIR 600
TPQIPWLLI LASLLALALI LAVCIAVNSR RRCGQKKLV INSGNGAVED RKPSGLNGEA 660
SKSQEMVHLV NKESSETPDQ FMTADETRNL QNVDMKIGV

```

Seq ID NO: 133 DNA sequence
Nucleic Acid Accession #: NM_002882
Coding sequence: 150-755

45
50
55
60

```

1 11 21 31 41 51
CGAGGTTCCG GTGTGGGGC GGAGGGAAGA GCGGGGCGGC GGGAGGCGCC GGGCCAGAC 60
GCGGAGGGA GAGGCTACGA GTAGCCCGCG AGAGGCCGCG GAGCCAGCGA CGACCGACCC 120
AGCCGAGCGC CCGCCGCGCG CGCGCCCCCA TGGCGGCGCG CAAGGACACT CATGAGGACC 180
ATGATACTTC CACTGAGAAT ACAGACGAGT CCAACCATGA CCCTCAGTTT GAGCCATAG 240
TTTCTCTTCC TGAGCAAGAA ATTAAACAC TGAAGAAGA TGAAGAGGAA CTTTTTAAAA 300
TGGCGGCAAA ACTGTTCCGA TTTGCCTCTG AGAACGATCT CCCGAAATGG AAGGAGCGAG 360
GCACTGCTGA CGTCAAGCTC CTGAAGCACA AGGAGAAAGG GGCATCCGC CTCCTCATGC 420
GGAGGACAAA GACCCCTGAG ATCTGTGCCA ACCACTACAT CACGCCGATG ATGGAGCTGA 480
AGCCCAAGCG AGGTAGOGAC CGTGCCCTGG TCTGGAACAC CCACGCTGAC TTGCGCGAGC 540
AGTGCCCCAA GCCAGAGCTG CTGCCCATCC GCTTCTCTGA TGCTGAGAA TGCACAGAAAT 600
TCAAAACAAA GTTTGAAGAA TGCAGGAAAG AGATCGAAGA GAGAGAAAAG AAAGCAGGAT 660
CAGGCAAAA TGATCATGCC GAAAAAGTGG CGGAAAAGCT AGAAGCTCTC TCGGTGAAGG 720
AGGAGACCAA AGGAGATGCT GAGGAGAAGC AATAAATCGT CTTATTTTAT TTTCTTTTCC 780
TCTCTTTCTT TCTCTTTT TAAAAAATT TACCTGCCCC CTCTTTTTCG GTTTGTTTTT 840
ATCTTTTCTT TTTTACAGG GAQTTTATAT AAAGAATCTA ACTC

```

Seq ID NO: 134 Protein sequence:
Protein Accession #: NP_002873

65
70

```

1 11 21 31 41 51
MAAAKTHED HDTSTENTDE SNHDPQFEPV VSLPEQBIKT LEEDEEELFK MRAKLFRPAS 60
ENDLPEWKER GTGDVKLLKH KERGAIRLLM RRDKTLKICA NHYITPMML KPNAGSDRAW 120
VMNTHADPAD ECPKPELLAI RFLNAENAQK FKTKFECRCK EIEEREKKG SGKNDHAQKV 180
AEKLEALS VK EETKEDAEK Q

```

Seq ID NO: 135 DNA sequence
Nucleic Acid Accession #: NM_000077.2
Coding sequence: 277-742

75
80
85

```

1 11 21 31 41 51
CCCAACCTGG GCGACTTCA GGTGTGCCAC ATTGCTAAG TGCTCGGAGT TAATAGCACC 60
TCTCTCGAGC ACTCGCTCAC GGCCTCCCTT TGCTGGAAGA GATACCGCG TCCCTCCAGA 120
GGATTGAGG GACAGGGTCG GAGGGGGCTC TTCGCGCAGC ACCGAGGAA GAAAGAGGAG 180
GGGCTGGCTG GTCACCAGAG GGTGGGGCGG ACCGCGTGGC CTGCGCGGCT GCGGAGAGGG 240
GGAGAGCAGG CAGCGGGCGG CGGGGAGCAG CATGGAGCCG GCGGCGGGGA GCAGCATGGA 300
GCCTTCGGCT GACTGCTGG CCAAGCGCCG GCGCGGGGT CCGGTAGAGG AGGTGCGGGC 360
GCTGCTGAGG GCGGGGGCGC TGCCCAACGC ACCGAATAGT TACGCTCGGA GGCAGATCCA 420
GGTCATGATG ATGGGAGCG CCGGAGTGGC GGAGCTGCTG CTGCTCCAGC GCGGAGAGCC 480

```

CAACTGGCCG GACCCCGCCA CTCTCACCGC ACCCGTGCAC GACGCTGCCG GGGAGGGCTT 540
 CCTGTGACAC CTGTGTGTGC TGCACCGGGC CGGGGCGCGG CTGSACGTGC GCGATGCTGT 600
 GGGCGGTCTG CCGGTGGACC TGGCTGAGGA GCTGGGCCAT CGCGATGTGC CACGGTACCT 660
 GGGCGCGCTG CGGGGGGCGA CCAGAGGCGA TAACCATGCC CGCATAGATG CCGCGGAAGG 720
 TCCTCAGACG ATCCCGGATT GAAAGAACCA GAGAGGCTCT GAGAAACCTC GGGAAACTTA 780
 GATCATCAGT CACCGAAGGT CCTACAGGGC CACAAC TGCC CCGCCACAA CCCACCCGCG 840
 TTTCTGAGTT TTTATTAGA AATAGAGCT TTTAAAAATG TCCTGCCTTT TAACGTAGAT 900
 ATATGCCTTC CCCCCTACC GTAAATGTCC ATTTATATCA TTTTATATAT ATTCTTATAA 960
 AATATGAAAA AGAGAAAAA CCGCTTCTGC CTTTCTACTG TGTGTGAGTT TTCTGGAGTG 1020
 AGCACTCAGC CCTTAGCGGC ACATTCATGT GGGCATTTCT TGCGAGCCTC GCAGCCTCCG 1080
 GAAGCTGTGC ACTTCATGAC AAGCATTTTG TGAAGTAGGG AAGCTCAGGG GGGTTACTGG 1140
 CTCTCTTGA GTCACTGC TAGCAAATGG CAGAACCAAA GCTCAATAA AATAAAAAA 1200
 ATTTTCATTC ATTCATC

Seq ID NO: 136 Protein sequence:
 Protein Accession #: NP_000068.1

1 11 21 31 41 51
 MEPAAGSSME PSADWLATRA ARGEVEEVRA LLEAGALPNA PMSYGRRPITQ VMMGSARVA 60
 ELLLLHGAEP NCADPATLTR PVHDAAREGF LDTLVVLHRA GARLDVRDAW GRLPVDLAEE 120
 LGHRDVARYL RAAAGSTRGS NHARIDAAEG PSDIPD

Seq ID NO: 137 DNA sequence
 Nucleic Acid Accession #: NM_058196.1
 Coding sequence: 104-421

1 11 21 31 41 51
 TGTGTGGGGG TCTGCTTGCC GGTGAGGGGG CTCTACACAA GCTTCCTTTC GGTCTATGCC 60
 GCGCCACCCC TGGCTCTGAC CATTCGTGTC TCTCTGGCAG GTCATGATGA TGGGCAGCGC 120
 CGAGTGTGCG GAGCTGTGTC TGCTCCACGG CGCGGAGCCC AACTGCGCGG ACCCGGCCAC 180
 TCTCACCCGA CCGGTGCACG ACCTGCCCCG GGAGGGCTTC CTGGACACGC TGGTGTGTCT 240
 GCACCGGGCC GGGGCGCGGC TGGACGTGCG CGATGCCTGG GCGCGTCTGC CCGTGGACCT 300
 GGCTGAGGAG CTGGGCCATC GCGATGTGCG ACCTGACCTG CGCGCGGCTG CCGGGGGCAC 360
 CAGAGGCACT AACCATGCCG GCATAGATGC CGCGGAAGGT CCCTCAGACA TCCCGGATTG 420
 AAAGAACCAG AGAGGCTCTG AGAAACCTCG GGAAACTTAG ATCATCAGTC ACCGAAGGTC 480
 CTACAGGGCC ACAACTGCCG CGGCCACAAC CCACCCCGCT TCGTAGTTT TCATTAGAA 540
 AATAGAGCTT TAAAAATGT CCTGCCTTTT AACGTAGATA TAAGCCTTCC CCCACTACCG 600
 TAAATGCCA TTTATATCAT TTTTATATA TTCTTATAA AATGTAAAA AGAAAAACAC 660
 CGCTTCTGCC TTTTCTACTG GTTGGAGTTT TCTGGAGTGA GCACTCAAGC CTAAGCGCA 720
 CATTCATGTG GGCATTTCTT GCGAGCCTCG CAGCCTCGG AAGCTGTGCA CTTATGACA 780
 AGCATTTTGT GAAGTAGGGA AGCTCAGGGG GGTACTGCG TTCTCTTGG TCACTGCT 840
 AGCAATGGC AGAACCAAG CTCAAATAA AATAAAATA TTTTCATTCA TTCCTC

Seq ID NO: 138 Protein sequence:
 Protein Accession #: NP_478103.1

1 11 21 31 41 51
 MMSGSARVAE LLLLHGAEPN CADPATLTRP VHDAAREGFL DTLVVLHRAE ARLDVRDAWG 60
 RLFLDLAEEL GHRDVARYLR AAGGTRGSN HARIDAAEGP SDIPD

Seq ID NO: 139 DNA sequence
 Nucleic Acid Accession #: NM_058197.1
 Coding sequence: 272-684

1 11 21 31 41 51
 CCCAACCTGG GGGGACTTCA GGTGTGCCAC ATTGCTAAG TGCTCGGAGT TAATAGCACC 60
 TCCTCGAGC ACTCGCTCAC GCGTCCCTCT TGCTGGAAA GATACCGCGG TCCCTCCAGA 120
 GGATTGTAGG GACAGGGTGC GAGGGGGCTC TTCCGCCAGC ACCGAGGAA GAAAGAGGAG 180
 GGGCTGGCTG GTCAACAGAG GGTGGGGCGG ACOCGTGCGG CTGCGCGGCT GCGGAGAGGG 240
 GAGAGCAGG CAGCGGGCGG CGGGAGCAG CATGGAGCGG GCGCGGGGA GCAGCATGGA 300
 GCGCGCGCGG GGGAGCAGCA TGGAGCCTTC GGTGACTGG CTGGCCACGG CCGCGGCGCG 360
 GGTGCGGTA GAGGAGGTGC GGGCGCTGCT GGAGCGGGG GCGCTGCCCA ACGCACCGAA 420
 TAGTTACGGT CAGAGGCGGA TCCAGGTGGG TAGAAGTCT GCAGCGGAG CAGGGGATGG 480
 CCGGCGACTC TGGAGGACGA AGTTTGACAG GGAATTGGAA TCAGGTAGCG CTTGATTCT 540
 CCGGAAAAAG GGGAGGCTTC CTGGGAGTT TTCAGAGGG GTTTGTAATC ACAGACCTCC 600
 TCCTGGCGAC GCCTGGGGG CTGGGAAAC CAAGGAAGAG GAATGAGGAG CCACGCGCGT 660
 ACAGATCTCT CGAATGCTGA GAAGATCTGA AGGGGGGAAC ATATTGTAT TAGATGGAAG 720
 TCATGATGAT GGGCAGCGCC CGAGTGGCGG AGCTGCTGCT GCTCCACGGC GCGGAGCCCA 780
 ACTGCGCGGA CCGCGCACT CTCAACCGAC CCGTGCACGA CGCTGCCCGG GAGGGCTTCC 840
 TGACACCGCT GGTGGTGTCT CACCGGGCGG GGGCGCGGCT GGAAGTGGC GATGCTGGG 900
 GCGCTCTGCC CGTGACCTG GCTGAGGAGC TGGGCCATCG CGATGTGCA CGGTACTGTC 960
 GCGCGCTGCG GGGGGGCACC AGAGGCAGTA ACCATGCCCC CATAGATGCC GCGGAAGGTC 1020
 CCTCAGACAT CCGCGATTGA AAGAACCAAG GAGGCTCTGA GAAACCTCGG GAACTTAGAT 1080
 CATCAGTCA CAGAGGTCTC ACAGGGCCAC AACTGCCCCC GCCACAACCC ACCCGCTTCT 1140
 CGTAGTTTTC ATTTAGAAAA TAGAGCTTTT AAAAAATGTC TGCCCTTTAA CGTAGATATA 1200
 TGCTTCCGCC CACTACCGTA AATGTCCATT TATATCATTT TTTATATATT CTTATAAAAA 1260
 TGTAAAAAAG AAAAACACCG CTCTGCTGCT TCTACTGTGT TGGAGTTTTC TGGAGTGAGC 1320
 ACTCAGCGCC TAGCGCACA TTCATGTGGG CATTTCTTGC GAGCCTCGCA GCCTCCGSA 1380
 GCTGTGACTC TCATGACAAG CATTTGTGTA ACTAGGGGAG CTCAGGGGGG TTAGTGGCTT 1440
 CTCTTGAGTC AACTGCTAG CAAATGGCAG AACCAGAGCT CAAATAAAAA TAAATAATT 1500

Seq ID NO: 140 Protein sequence:
Protein Accession #: NP_478104.1

5
1 11 21 31 41 51
MEPAAGSSME PAAGSSMEPS ADWLATAAAR GRVEEVRALL EAGALPNAPN SYGRRPIQVG 60
10 RRSAAAGAGDG GRLWRTKFAE ELESSESASIL RKKGRPLPEF SEGVCNHRPP PGDALGAWET 120
KEEE

Seq ID NO: 141 DNA sequence
Nucleic Acid Accession #: NM_058195.1
Coding sequence: 163-684

15
1 11 21 31 41 51
CCTCCCTACG GCGCGCTCCG GCAGCCCTTC CCGCGTGGGC AGGGCTCAGA GCGTTCOGA 60
20 GATCTTGGAG GTCCGGGTGG GAGTGGGGGT TGGGGTGAA GGTGGGGGCG 120
GGGCGCGCTC AGGGAAGGCG GGTGGCGGCC TGGCGGGCGG AGATGGGCGG GGGCGGGTGC 180
GTGGGTCCCA GTCTGCAGTT AAGGGGCGAG GAGTGGCGCT GCTCCTCTCT GGTGCCAAAG 240
GGCGGGCGCG GCGGTGCGCA GCTCGGCCCT GGAGGCGGCG AGAACATGGT GCGCAGGTTT 300
TTGGTGACCC TCCGATTTCG GCGCGCGTGC GGCGCGCGCG GAGTGAGGGT TTTCTGTGTT 360
CACATCCCGC GGCTCAGGGG GGAGTGGGCA GCGCCAGGGG CGCCCGCGCG TGTGGCCCTC 420
25 GTGCTGATGC TACTGAGGAG CCAGCGTCTA GGGCAGCAGC CGCTTCTCTAG AAGACCAGGT 480
CATGATGATG GGCACGCGCC GAGTGGCGGA GCTGCTGCTG CTCCACGGCG CGGAGCCCAA 540
CTGGCGCGAC CCGCCACTC TCACCGGACC CTTGCACGAC GCTGCCCGGG AGGGCTTCCT 600
GGACACGCTG GTGGTGTGCG ACCGGGCGCG GCGCGCGCTG GACGTGCGCG ATGCGCTGGG 660
CGTCTGCGCC GACTACCGGT CTGAGGAGCT GGGCCATGCG GATGTGCGAC GGTACCTGGG 720
30 CGCGCGTGGC GGGGCGACCA GAGGCAGTAA CCGTCCCGCG ATAGATGCGG CGGAAGGTCC 780
CTCAGACATC CCGGATTGAA AGAACCGAG AGGCTCTGAG AAACTCTCGG AAACCTAGAT 840
CATCAGTCCG CGAAGGTCCT ACAGGGCCAC AACTGCCCGG GCCACAACCC ACCCGCTTTT 900
CGTAGTTTTC ATTTAGAAAA TAGAGCTTTT AAAAATGTCC TGCCCTTTAA CGTAGATATA 960
TGCTCTCCCG CACTACCGTA AATGTCCATT TATATCATTT TTTATATATT CTTATAAAAA 1020
35 TGTAAAAAAG AAAAACACCG CTTCTGCCTT TTCCTGTGTT TGGAGTTTTC TGGAGTGAGC 1080
ACTCAGCGCC TAAGCGCACA TTCTGTGTGG CATTTCTTGC GAGCCTCGCA GCCTCCGGAA 1140
GCTGTGACT CTATGACCAAG CATTTTGTGA ACTAGGGAAG CTCAGGGGGG TTACTGGCTT 1200
CTCTTGATG ACATGTCTAG CAAATGGCAG AACCAAGACT CAAATAAAAA TAAATAAATT 1260
40 TTCATTCATT CACTC

Seq ID NO: 142 Protein sequence:
Protein Accession #: NP_478102.1

45
1 11 21 31 41 51
MGRGRCVGPS LQLRGQEWRC SPLVFKGGAA AELGPGGEE NMVRRFLVTL RRRACGPPR 60
VRVFFVHIPR LTGEWAAPGA PAAVALVLM LRSRLGQQP LPRRPGHDDG QRPSSGAAAA 120
50 PRGALQLRRP RHSHPTARR CPGLPGHAG GAAPGRGAAG RARCLGPSAR GPG

Seq ID NO: 143 DNA sequence
Nucleic Acid Accession #: NM_018131
Coding sequence: 412..1107

55
1 11 21 31 41 51
GAAATTGCAC ACTTAAAGAC ATCAGTGGAT GAAATCACAA GTGGGAAAGG AAAGCTGACT 60
60 GATAAAGAGA GACAGAGACT TTTGGAGAAA ATTGAGTCC TTGAGGCTGA GAAGGAGAAG 120
AATGCTTATC AACTCACAGA GAAGGACAAA GAAATACAGC GACTGAGAGA CCAACTGAAG 180
GCCAGATATA GTACTACCGC ATTGCTTGAA CAGCTGGGAG AGACACGAG AGAAGGAGAA 240
AGGAGGGAGC AGGTGTGAA AGCCTTATCT GAAGAGAAAG ACCTATTGAA ACAACAGTTG 300
TCTGCTGCAA CCTCAGAAAT TGCTGAACCTT GAAAGCAAAA CCAATACACT CCGTTTATCA 360
CAGACTGTGG CTCCAAACTG CTTCAACTCA TCAATAAATA ATATTATGTA AATGGAAATA 420
65 CAGCTGAAAG ATGCTCTGGA GAAAAATCAG CAGTGGCTCG TGTATGATCA GCAGCGGGAA 480
GTCTATGTAA AAGGACTTTT AGCAAAATGC TTTGAGTTGG AAAAGAAAAC GGAACAGCT 540
GCTCAITCAC TCCACAGCA GACAAAAAAG CCTGAATCAG AAGGTTATCT TCAAGAAGAG 600
AAGCAGAAAT GTTACAGCA TCTCTGGCA AGTGCAAAAA AAGATCTTGA GGTGAAACGA 660
70 CAAACCATAA CTCAGCTGAG TTTTGAAGT AGTGAATTTT GAAGAAAATA TGAAGAAACC 720
CAAAAAGAA TTCACTATT AAATCAGCTG TTGTATTAC AAAGAAGGGC AGATGTGCAA 780
CATCTGGAAG ATGATAGGCA TAAACAGAG AAGATACAAA AACTCAGGGA AGAAGATGAT 840
ATTGCTAGGG GAAAACTTGA AGAAGAGAAG AAGAGATCCG AAGAGCTCTT ATCTCAGGTC 900
CAGTCTCTTT ACACATCTCT GCTAAAGCAG CAAGAAGAAC AAACAGGGT AGCTCTGTG 960
75 GAACAACAGA TGCGGCGATG TACTTTAGAC TTTGAAAATG AAAAAGTCTG CCGTCAACAT 1020
GTGCGAGCAT AATTGCTATG AATTCTTAAG GAGCTCCGAA AAGCAAGAAA AAATAACACA 1080
GTTGGAATCC TTGAAACAGC TTCTAGAGTT TGCCATCACA GAGCCATTAG TCACTTTCCA 1140
AGGAGAGATT GAAACAGAG AAAAAAGTTG CGCCTCACCA AAAAGTCCCA CTGCTGCACT 1200
CAATGGAAGC CTGGTGAAT GTCCCAAGTG CAATATACAG TATCCAGCCA CTGAGCATCG 1260
CGATCTGCTT GTCCATGTGG AATCTGTTT AAAGTAGCAA AATAAGTATT TGTTTTGATA 1320
80 TTAAGAAGAT CAATAGTGA TTTCTGTGA GCTGTGTGGC ATTTGAATT ATATATTCCA 1380
CATTTGTCAT AAAAGTCCCT ATCTACCTTT GACACTCCAG CATGCTAGTG AATCATGTAT 1440
CTTTTAGGCT GCTGTGCAAT TCTCTGGCA GTGATACCTC CCGTACATGG TTATCATCA 1500
GGCTGCAATG ACAGAAATGT GTGAGCAGCG TCTACTGAGA TACTAACATT TTGCACTGTC 1560
85 AAAATACTTG GTGAGGAAAA GATAGCTCAG GTTATTGCTA ATGGGTAAAT GCACCAGCAA 1620
GCAAAATATT TTATGTTTCG GGGGTTTGA AAAATCAAG ATAATTAACC AAGGATCTTA 1680
ACTGTGTTG CATTTTTCAT CCAAGCACTT AGAAAACCTA CAATCCTAAT TTTGATGTCC 1740
ATTGTTAAGA GGTGGTGATA GATACATTTT TTTTTCATA TTGTATAGCG GTTATTAGAA 1800

5 AAGTTGGGGA TTTCTTGAT CTTTATTGCT GCTTACCATT GAACTTAAC CCAGCTGTGT 1860
TCCCCAACTC TGTCTGCGC ACGAAACAGT ATCTGTTTGA GGCATAATCT TAAGTGGCCA 1920
CACACAATGT TTTCTCTTAT GTTATCTGGC AGTAACTGTA ACTTGAATTA CATTAGCACA 1980
TTCTGCTTAG CTAAATTTGT TAAATAAAC CCAATGTAGCC CTCTCATTGT 2040
ATTGACAGTA TTTTAGTTAT TTTTGGCATT CTAAAGCTG GGCATGTAA TGATCAGATC 2100
TTTGTTTGTC TGAACAGGTA TTTTATACA TGCTTTTGT AAACCAAAA CTTTAAAT 2160
TCTTCAGGTT TTCTAACATG CTTACCACTG GGCTACTGTA AATGAGAAAA GAATAAAAT 2220
ATTTAATGTT TT

10

Seq ID NO: 144 Protein sequence:
Protein Accession #: NP_060601

15

1 11 21 31 41 51
MEIQLKDALE KNQQLVYDQ QREVVYKGLL AKIFELEKKT ETAHSLPQQ TKKPESEGYL 60
QEEKQKCYND LLASAKKDL VERQTITQLS FELSEPRRY ESTQREVNHL NQLLYSQRR 120
DVQHLEDDRH KTEKIQLRE ENDIARGKLE EEKRSSEELL SQVQSLYTSL LKQQEEQTRV 180
20 ALLEQQMQAC TLDPENEKLD RQHVQHQLHV ILKELRKARK NNTVGILETA S

25

Seq ID NO: 145 DNA sequence
Nucleic Acid Accession #: NM_001168
Coding sequence: 50..478

30

1 11 21 31 41 51
CCGCCAGATT TGAATCGCGG GACCGTGTGG CAGAGGTGGC GCGCGCGGCA TGGGTGCCCC 60
GAGTGTGCCC CCGCTCTGCG AGCCCTTTCT CAAGGACCAC CGCATCTCTA CATTCAAGAA 120
CTGGCCCTTC TTGAGGGGCT GCGCCTGCAC CCGGAGCGG ATGGCCGAGG CTGGCTTCAT 180
CCACTGCCCC ACTGAGAAGC AGCCAGACTT GCGCCAGTGT TTCTTCTGCT TCAAGGAGCT 240
GGAGGCTGG GAGCCAGATG ACGACCCCAT AGAGGAACAT AAAAAGCATT CGTCCGGTTG 300
CGCTTTCTCT TCTGTCAAGA AGCAGTTTGA AGAATTAACC CTGGTGAAT TTTTGAAACT 360
35 GGACAGAGAA AGAGCCAAGA ACAAATTCG AAAGGAAACC AACATAAGA AGAAAGAATT 420
TGAGGAAACT CGCAAGAAAG TGCGCCGTGC CATCGAGCAG CTGGCTGCCA TGGATTGAGG 480
CCTTGGCCCG GAGCTGCCTG GTCCAGAGT GGCTGCACCA CTCCAGGGT TTATTCCCTG 540
GTGCCACCAG CCTTCTGTG GGCCTCTTAG CAATGTCTTA GGAAGGAGA TCAACATTTT 600
CAAATTAGAT GTTTCAACTG TGCTCCTGTT TTGTCTTGAA AGTGGCACA GAGGTGCTTC 660
40 TGCTGTGCA GCGGGTGTG CTGGTAACAG TGGCTGCTTC TCTCTCTCTC TCTCTTTTTT 720
GGGGGCTCAT TTTTGTGTTT TTGATTCCTG GGTCTACCA GTGAGAAAGT AGGGAGGAAG 780
AAGGCASTGT CCGTTTGTCT AGAGCTGACA GCTTTGTTCG CCGGGGAGAG GCGTTCCACA 840
GTGAATGTGT CTGGACCTCA TGTGTGTGAG GCTGTACAG TCGTGTGAGT GGAAGTGGCA 900
GGTGCCTGTT GAATCTGAGC TGCAGGTTC TTATCTGTCA CACTGTGCC TCCTCAGAGG 960
45 ACAGTTTTTT TGTGTGTGTT TTTTGTGTT TTTTGTGTT GGTAGATGCA TGACTTGTGT 1020
GTGATGAGAG AATGAGAGCA GAGTCCCTGG CTCTCTACT GTTTAAACAC ATGGCTTTCT 1080
TATTTGTGTT GAATGTGTTA TTCACAGAA AGCACAAACT ACAATTAATA CTAAGCACAA 1140
AGCCATTCTA AGTCATTGGG GAAACGGGGT GAACTTCAGG TGGATGAGGA GACAGAAATG 1200
AGTGATAGGA AGCGTCTGCG AGATACTCTT TTTGCCACTG CTGTGTGATT AGACAGGCC 1260
50 AGTGAGCCCG GGGGCACATG CTGGCCGCTC CTCCCTCAGA AAAAGGCAGT GGCCTAAATC 1320
CTTTTAAAT GACTTGGCTC GATGCTGTGG GGGACTGGCT GGGCTGTGCG AGGCCGTGTG 1380
TCTGTACGCC CAACCTTCAC ATCTGTCAAG TTCTCCACAC GGGGGAGAGA CGCAGTCCG 1440
CCAGTCCCGC GCTTCTTTG GAGGCAGCAG CTCCCGCAGG GCTGAAGTCT GGCCTAAGAT 1500
55 GATGGATTGG ATTCGCCCTC CTCCCTGTCA TAGAGCTGCA GGGTGGATTG TTACAGCTTC 1560
GCTGGAAACC TCTGGAGGTC ATCTCGGCTG TTCCTGAGAA ATAAAAAGCC TGTCAATTC

60

Seq ID NO: 146 Protein sequence:
Protein Accession #: NP_001159

65

1 11 21 31 41 51
MGAPTLPFAW QPFLKDHRS TFKWPPLEG CACTPERMAE AGFIHCPTEN EPDLAQCFPC 60
FKELEGWEPD DDPIEEHKH SSGCAFLSVK KQFELTLGE FLKLDREARAK NRIAKETNNK 120
KKEFETAARK VRRRIEQLAA MD

70

Seq ID NO: 147 DNA sequence
Nucleic Acid Accession #: NM_014176.1
Coding sequence: 127-720

75

1 11 21 31 41 51
GCGCGCAGCG CTGGTACCCC GTTGGTCCGC GCGTGTGCTG GTTGTGAGGG GTGTGAGCTC 60
AGTGCATCCC AGGCAGCTCT TAGTGTGGAG CAGTGAACCT TGTGTGTTTC CTCTACTCTG 120
GGGATCATGC AGAGAGCTTC ACGTCTGAAG AGAGAGCTGC ACATGTTAGC CACAGAGCCA 180
CCCCAGGCA TCACATGTTG GCAAGATAAA GACCAATGAG ATGACCTGCG AGCTCAAATA 240
TTAGGTGGAG CCAACACACC TTATGAGAAA GGTGTTTTTA AGCTAGAAAT TATCATTCCT 300
GAGAGGTACC CATTGAACC TCCTCAGATC CGATTCTCA CTCCAATTTA TCATCCAAC 360
ATTGATTCTG CTGGAAGGAT TTGTCTGGAT GTTCTCAAAAT TGCCACCAAA AGGTGCTTTG 420
80 AGACCATCCC TCAACATGCG AACTGTGTTG ACCTCTATTC AGCTGCTCAT GTCAGAACCC 480
AACCTGATG ACCGCTCAT GGCTGACATA TCCTCAGAAAT TTAATATAA TAAGCCAGCC 540
TTCCTCAGA ATGCCAGACA GTGCACAGAG AAGCATGCAA GACAGAAACA AAAGCTGAT 600
GAGGAAGAGA TGCTTGATAA TCTACCAGAG GCTGGTGACT CCAGAGTACA CAACTCAACA 660
CAGAAAGGA AGGCCAGTCA GCTAGTAGGC ATAGAAAGA AATTTCTATC TGATGTTTAG 720
85 GGGACTTGTG CTGCTTCATC TTAGTTAATG TGTCTTTGCG CAAGGTGATC TAAGTTGCTT 780
ACCTGAATTT TTTTATAA TATATTGAT GACATAATTT TTGTGTAGTT TATTTATCTT 840
GTACATATGT ATTTTGAAT CTTTAAACC TGAAAAATA ATAGTCATTT AATGTTGAAA 900

Seq ID NO: 148 Protein sequence:
 Protein Accession #: NP_054895.1

5
 1 11 21 31 41 51
 10 MQRASRLKRE LHMLATEPPP GITCWQDKDQ MDDLRAQILG GANTPYEKGV FKLEVIIPER 60
 YPFEPPIQIR LTPYIYHNID SAGRICLDVL KLPPKGAWRP SLNIATVLTS IQLLMSEFNP 120
 DDPLMADISS EFKYNKPAFL KNARQWTEKH ARQKQKADEE EMLDNLPEAG DSRVHNSTQK 180
 RKASQLVGIE KKFHPDV

Seq ID NO: 149 DNA sequence
 Nucleic Acid Accession #: NM_003812
 Coding sequence: 224-2722

15
 1 11 21 31 41 51
 20 TCCTCTGCGT CCCGCCCGCG GAGTGGCTGC GAGGCTAGGC GAGCCGGGAA AGGGGGCGCC 60
 GCCAGCGCCC GAGCCCGCGG CCCCGTCCCC CGAGCCCGGA GCCCCTGCC CGCGGCGGCA 120
 CCATGCGCGC CGAGCCCGCG TGACCGGCTC CGCCCGCGGC CGCCCGCAG CTAGCCCGGC 180
 GCTCTCGCGG GCCACACGGA GCGGCGCCCG GGAGCTATGA GCCATGAAGC CGCCCGGCAG 240
 CAGCTCGCGC CAGCCGCCCC TGGCGGGCTG CAGCCTTGCC GGCGCTTCCT CGCGCCCCCA 300
 ACGCGGCCCG CGCGGCTCGG TGCCCTGCCAG CGCCCGGGCC CGCAGCGCGC CCTGCGCGCT 360
 25 GCTTCTCGTC CTTCTCTGCG TGCCCTCGCT CGCCGCGCTG TCCCGGCCCC GCGCTCGGG 420
 GGCTGCTGCG CCCAGCGCTC CGCATTGGAA TGAAACTGCA GAAAAAATT TGGGAGTCCT 480
 GGAGATGAA GACAAATACAT TGCAACAGAA TAGCAGCAGT AATATCAGTT ACAGCAATGC 540
 AATGACAGAA GAAATCACAC TGCCTTCAAG ACTCATATAT TACATCAACC AAGACTCGGA 600
 AAGCCCTTAT CAGCTTCTTG ACACAAAGGC AAGACACCAG CAAAAACATA ATAAGGCTGT 660
 30 CCATCTGCGC CAGGCAAGCT TCCAGATTGA AGCCTTCGGC TCCAAATTC TCTTGAAGCT 720
 CATACTGAAC AATGGTTTGT TGTCTTCTGA TTAATGAGG ATTCACTACG AAAATGGGAA 780
 ACCACAGTAC TCTAAGGGTG GAGAGCAGTG TTAATACCAT GGAAGCATCA GAGGCGTCAA 840
 AGACTCCAAG GTGGCTCTGT CAACCTGCAA TGGACTTCAT GGCAATGTTG AAGATGATAC 900
 CTTGCTGTAT ATGATAGAGC CACTAGAGCT GGTTCATGAT GAGAAAAGCA CAGGTGAGCC 960
 35 ACATATAATC CAGAAAACCT TGGCAGGACA GTATTCTAAG CAAATGAAGA ATCTCACTAT 1020
 GGAAGAGAGT GACCAAGTGC CCTTCTCTCT TGAATTACAG TGGTTGAAAA GAAGGAAGAG 1080
 AGCAGTGAAT CCATCACGCT GTATATTGTA AGAATGAAA TATTGGAAC TTATGATTGT 1140
 TAATGATCAC AAAACGTATA AGAAGCATCG CTCTTCTCAT GCACATACCA ACAACTTTGC 1200
 AAAGTCGCGT GTCACCTGTG TGGATTCTAT TTACAAGGAG CAGCTCAACA CCAGGTTTGT 1260
 40 CCTGGTGGCT GTAGAGACCT GGACTGAGAA GGATCAGATT GACATCACC CCAACCTGT 1320
 GCAGATGCTC CATGAGTTCT CAAATATCCG GCAGCGCATT AAGCAGCATG CTGATGCTGT 1380
 GCACCTCATC TCGCGGGTGA CATTTCACTA TAAGAGAAGC AGTCTGAGTT ACTTTGAGG 1440
 TGTCTGTTCT CGCACAGAGC GAGTTGGTGT GAATGAGTAT GGTCTTCAA TGGCAGTGGC 1500
 45 ACAAGTATTA TCGCAGAGCC TGGCTCAAAA CCTTGAATC CAATGGGAAC CTTCTAGCAG 1560
 AAAGCCAAAA TGTGACTGCA CAGAACTCTG GGGTGGCTGC ATCATGGAGG AAACAGGGGT 1620
 GTCCCATTTT CGAAATATTT CAAAGTGCAG CATTTTGGAG TATAGAGACT TTTTACAGAG 1680
 AGGAGGTGAA GCTGCGCTTG TCAACAGGCC AACAAAGCTA TTTGAGCCCA CGGAATGTGG 1740
 AAATGGATAG GTGGAAGCTG GGGAGGAGTG TGAATGTGGT TTTCAATGTT AATGCTATGG 1800
 50 ATTAGTCTGT AAGAAATGTT CCTCTCCAA CGGGGCTCAC TGCAGCGAG CGCCCTGCTG 1860
 TAACAATACC TCAATGCTTT TTCAGCCACG AGGGTATGAA TGCCGGGATG CTGTGAACGA 1920
 GTGTGATATT ACTGAATATT GTACTGGAGA CTCTGGTCAG TGCCCAACCA ATCTTCATAA 1980
 GCAAGACGGA TATGCAATGA ATCAAAATCA GGGCGCGTGC TACAATGGCG AGTGCAAGAC 2040
 CAGAGACAC CAGTGTGAGT ACATCTGGGG AACAAAGGCT GCAGGCTCTG ACAAGTTCTG 2100
 55 CTATGAAAA CTGAATACAG AAGGCACTGA GAAGGGAAGG TGCGGGAAGG ATGAGAGCCG 2160
 GTGGATTAGT TGCAGCAAA ATGATGTGTT CTGTGGATTC TTAATCTGTA CCAATCTTAC 2220
 TCGAGCTCCA CGTATTGGCT AACTTCAGGG TGAGATCATT CCACTTCTCT TCTACCATCA 2280
 AGGCGGGGTG ATGACTGCA GTGGTGCCCA TGTAGTTTGA GATGATGATA CGGATGTGGG 2340
 CTATGTAGAA GATGGAACCG CATGTGCCCC GTCTATGATG TGTTAGATC GGAAGTGCCT 2400
 60 ACAAATTCAA GCCCTAAATA TGAGCAGCTG TCCACTCGAT TCCAAGGGTA AAGTCTGTTT 2460
 GGGCCATAGG GTGTGTAGTA ATGAAGCCAC TGCATTTGT GATTTACCT GGGCAGGGAC 2520
 AGATTGCAAT ATCCGGGATC CAGTTAGGAA CCTTCACCCC CCAAGGATG AAGGACCCAA 2580
 GGGTCTTAGT GCCACCAAT TCATAATAGG CTCCATCGCT GGTGCCATCC TGGTAGCAGC 2640
 65 TATTGTCTCT GGGGGCACAG GCTGGGGATT TAAAAATGTC AAGAGAGAA GGTTCGATCC 2700
 TACTCAGCAA GGGCCCATCT GAATCAGCTG CGCTGGATGG ACACCGCCTT GCATCTGTTG 2760
 ATTTCTGGTA TGACATACTC GCAGCAGTGT TACTGGAAC ATTAAGTTTG TAAACAAAAC 2820
 CTTTGGGTGG TAATGACTAC GGAGCTAAAG TTGGGGTGAC AAGGATGGGG TAAAGAAAA 2880
 CTGTCTCTTT TGGAAATAAT GTCAAAGAAC ACCTTTCACC ACCGTGAGT AAACGGGGGA 2940
 GGGGGCAAAA GACCATGCTA TAAAAAGAAC TGTTCAGAA TCTTTTTTT TCCCTAATGG 3000
 70 ACGAAGGAAC AACACACACA CAAAAATTAA ATGCAATAAA GGAATCATT AAAAA

Seq ID NO: 150 Protein sequence:
 Protein Accession #: NP_003803

75
 1 11 21 31 41 51
 80 MKPPGSSSRQ PPLAGCSLAG ASCGPQRGPA GSVPASAPAR TPPCRLLLV LLLPPLAASS 60
 RPRAGAAAP SAPHWNETA KNLGVLADED NTLQNNSSSN ISYSNAMQKE ITPSLRIYY 120
 INQDSBPYH VLDTKARHQ KHNKAVHLAQ ASPQIEAFGS KFIIDLILNN GLLSSDYVEI 180
 HYENGKPYYS KGGHECYTHG SIRGVKDSKV ALSTCNGLHG MPEDDTFVYM IEPLRLVDE 240
 KSTGRPHIQ KTLAQYYSKQ MKNLTMERGD QWPLSELQW LKRRKRAVNP SRGIFEMKY 300
 LELMIVNDK TYKHSRSHA HTNNFAKSV NLVDSIYKEQ LNRVVLVAV ETWTEKDQID 360
 ITTNPVQMLH EFSKYRQRIK QHADAVHLIS RVTFRYKRSS LSYFGVCSR TRGVGVNEYG 420
 85 LPMVAQVLS QSLAQLGLQ WEPSSRKPKC DCTESWGGCI MEETGVSHSR KFSKCSILEY 480
 RDLQRGGGA CLFNRPTKLF EPTCEGNGYV EAGEECDCGF HVBCEYLCKC KCSLNGAHC 540
 SDGPCCNNTS CLFPPRGYEC RDAVNECDIT EYCTGDSGQC PPNLEKQDGY ACNQNGRCY 600
 NGECKTRDQ CQYIWTGKAA GSDKFCYEKL NTEGTEKGNC GRDGRNWIQC SKHDVFCGFL 660

LCTNLTRAPR IGQLQGEIIP TSFYHQGRVI DCSGAHVVD DTDVGVYVED GTPCGFSMMC 720
 LDRKCLQIQA IAMSCLPLDS KGVKCSGHGV CSNEATCICD FTWAGTDCSI RDPVRNLHPP 780
 KDEGPKGPSA TNLIIIGSIAG AILVAAIVLG GTGWGFKNVK KRRFDPTQQG PI

5

Seq ID NO: 151 DNA sequence
 Nucleic Acid Accession #: NM_023915
 Coding sequence: 250-1326

10 1 11 21 31 41 51
 GGCACGAGGG TTTCGTTTTC ATGCITTACC AGAAAATCCA CTTCCTGCC GACCTTAGTT 60
 TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120
 GTGAATGGAC AGCCAGCCAC CACAAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180
 15 CCCACGCTC AATGCTCCC AAGTGTTCCT TGACACGCAT CTTTGCCTAC AGTGCAATCAC 240
 AACTGAAGAA TGGGGTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCAACGGC 300
 CAAGAGAGTC ACAATTCAGG CAACAGGAGC GACCGGCCAG GAAAGAACAC CACCCTTCAC 360
 AATGAATTTG ACACAATTGT CTTGCGGTG CTTTATCTCA TTATATTGT GGCAGCATC 420
 TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAC CAGCTTCATA 480
 20 TTCTATCTCA AAAACATAGT GGTTCGAGAC CTCATAATGA CGCTGACATT TCCATTTCGA 540
 ATAGTCCATG ATGACGAGT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600
 TCAGTTTTGT TTTATGCAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCATT 660
 GATCGCTATC TGAAGTGGT CAAGCCATTT GGGGACTCTC GGATGTACAG CATAACCTTC 720
 ACGAAGGTTT TATCTGTTG TGTTCGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780
 25 ATCCTGACAA ATGCTCAGC AACAGAGGAC AATATCCATG ACTGCTCAA ACTTAAAGT 840
 CCTTTGGGGG TCAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900
 GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960
 AGGCAATTCA TAAGTCAGT AAGCCGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT 1020
 GTGGCTGTG TTTTACCTG CTTTCTACCA TATCACTTGT GCAGAAITCC TTTTACTTTT 1080
 30 AGTCACCTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCATATTA CTGCAAGAA 1140
 ATTACACTTT TCTGTCTGCG GTGTAATGTT TGCTGGATC CAATAATTTA CTTTTCATG 1200
 TGTAGGTCAT TTCAAGAAG GCTGTTCAAA AAATCAAATA TCAGAACAG GAGTGAAAGC 1260
 ATCAGATCAC TGCAAAGTGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTACACTGAT 1320
 GTGTAGGCCT TTTATTGTTT GTTGAATCG ATATGTACAA AGTGTAATAA AATGTTCTT 1380
 35 TTCATTATCC TTAATAAAAA AA

Seq ID NO: 152 Protein sequence:
 Protein Accession #: NP_076404

40 1 11 21 31 41 51
 MGFNLTAKL PNNELHQES HNSGNRSDGP GKNTTLHNEF DTIVLPVLYL IIFVASILLN 60
 GLAVNIFFHI RNKTSFIFYL KNIVVADLIM TLTFPFRIVH DAGFGPWYFK FILCRYTSVL 120
 FYANMYTSIV FLGLISIDRY LKVVKPPGDS RMYSTPTFKV LSVVVMVMA VLSLPNIILT 180
 45 NQPTEDNTH DCSKLKSPGL VKWHTAVTYV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQF 240
 ISQSSRRRKH NQSRVUVAV PFTCFPLPYL CRIPFTPSHL DRLLDESAQK ILYYCKEITL 300
 FLSACNVCLD PIIFYPMCRS PSRLPKKSN IRTRESIRS LQSVRRSEVR IYYDYTDV

50

Seq ID NO: 153 DNA sequence
 Nucleic Acid Accession #: D80008.1
 Coding sequence: 149-739

55 1 11 21 31 41 51
 GTTCGGCGCC AAAGCGCGGA GCGGAGGCGG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60
 CGAAGAGAGT GAGGCGCGGA GAGCCAGAT ACCATTTTGG CGTGAGAGCT GGTGTTGGC 120
 AAGGCGCGCG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180
 CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCCTC AAGGAGGATG GACTCAGACA 240
 60 AGTTCTGGAG GAGATGAAGG CTTTGTATGA ACAAACCAG TCTGATGTA ATGAAGCAAA 300
 GTCAGGTGGA CGAAGTAGT TGATACCAAC TATCAAATT OGACACTGTT CTCTGTTAAG 360
 AAATCGACGC TGCACTGTAG CATACCTGTA TGACCGCTTG CTTGGATCA GAGCACTCAG 420
 ATGGGAATAT GGTAGCTCT TGCCAAATGC ATTACGATT CACATGGCTG CTGAAGAAAT 480
 GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTCAC TGGGAGGAGA 540
 65 TGAAGGTTTG GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGTCCG 600
 GTGTCTAAAA GACTATGGAG AATTGGAAGT TGATGATGGC ACTTCAGTCC TATTAAAAAA 660
 AAATAGCCAG CACTTTTAC CTGATGGAA ATGTGAGCAG CTGATCAGAC AAGGAGTCC 720
 GGAGCACATC CTGTATGAC CATGCGCGGA GGCATTCCA GGCTTCACTC AACTCATGGA 780
 CTCCTCTGTA CTCACCTCTC CCACCACTCC CTCACCTCC CTCTTGATT TTAGAAGCTA 840
 70 TAGACATTGT TTAAGATAAC TAAGAATACT TGGCTAAGAA GTATAATTG CTAACATTA 900
 AGGACTTTCT TTTTAAATG TTGTACACTA TTCTTCTTAC TCTTTTGGG TTTTGGTTT 960
 GTTTTGTAGA GACTGTCTCA CTATGTTGCC CAAGCTGGTC TCAAACCTCT GGCCTCAAGC 1020
 AGTCTCCCA CCTTAGCTTC TCAAAGTGT GAGATCACAG GGTGAGCCA CTGCACCCGG 1080
 CCCCTACTCC TTTTCTAAT AAGCTGTATC TGTAATCACA GCATTCTTAC AGTTGTTACA 1140
 75 GTGTGTTTT TAAATGAAG TAAACATGGT TACATTGAA TCTCTTAAAT AAGCAGTCAC 1200
 TTGGCTGAG AGGAAGAAGG TAGATCCTGT GTGTCTGTT TTCTGGTCAT GTGATTGTA 1260
 CAAGCTAGAG AGCTGAATTT CTGAGATACA CATTTTCAA TCACATGCAA GTGAAGATGA 1320
 TGGTCTGAG AAATTTTCAG TATATATAAT GTTTAATGAC ATACTAATTT ATCATCTGGC 1380
 TATTTGGGAA GGAAGGACAC ACATGGATTT TGCACATTTC CACCATGGTG GCTGGTGTGG 1440
 80 CTGTGGGCTA TGGGGTATC ACCAGATCA CCACITTTGA AGGGGACAGT GAAATGGGG 1500
 CTAGAGAAAG AACITTTGAT AGTTTCCCT GAGATTGAGA TTGACTGAAA AGTCACATGA 1560
 AGAGTGTAT GTCTTTTAAAT GGTATGTTTT AAACAGCTGA CATTTAAAT TTTGATGAAA 1620
 TCCAGTTTAT TCGTTTGTTC TTTTATGCTT TGGGTGTGAC ATCCGAGAAA TCTTTTCCCA 1680
 TCCCAAGATC ACAATTTTTT TTCTTTTATA CTCTAGAAAG TGTTATAATT TTAAGCTTTA 1740
 TACTTTGGCT TATGACCCGT TTTTTTTTT GTTTTGTGTT GTTTTTCGT TTTTTCCTT 1800
 85 GTTTTGAAT GAGGTCTTGT TCTGTACCC AGGCTGGGGT GCAGTGGCGT GATCTTGGCT 1860
 CACTGCAATC TCTATCCCT GGGTCAAGT GATTCTCTTG TCTCAGCTC CCAAGTAGCT 1920
 GGGATTACAG GCACAGGCCG CCAAGCCTGG CTAATTTTTG TATTTTGTAGT AGAGACAGAG 1980

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TTTTACCATG TTGGCCAGGC TGGTTTCAAA CTCCTGACCT CAAGTGACCC ACCTTGGCCT 2040
CCCAAAGTTT TGGGATTACA AGTGTGGGCC ACCGCGGCCA GCCTATGATC CATTTTGAAT 2100
GAATTTTTTA TAGTGTCCAA GGTGTCAATC CACCTTCACT TTTTCTTGGG AATATAGATA 2160
TCAGCTGTTT TCACATACCAT TTTTGAAGA GACTGCCCTT TGCTCTATCA CCTTGTGATT 2220
TTTGTAAAAA AGTAGTTGTC AATGTATATG TGGGTTTATT TCAGGACTCT GTTTTGTGTC 2280
ATTGACCTGT TTTTCTCTCC TGAATGCCAA TACCATATTT GTATGTAGTG TATGTAATTT 2340
TCTAATAATT CTGAAACAGC ATAGTATTAA TGTGTATAT TTTTGTCTGT GTTTGTATTT 2400
TTTGTAGAGA TGGGGTTTCA CCGTGTGGGC CAGGCTGTGT TGAACCTCTG AGCTAAAGCA 2460
ATACACTTGC CTGCTCTCTC CCATGTGCTG GGATTACAGG CGTGAGCCTT GGTGCTGGCC 2520
CAGTGTACCA CATTTCTTTT TGAGATTGTG TTTGGCTATG TTAAGTCCCT TGCTTTTGAT 2580
GTGAAATTTG GGAACAGGCA GGGTGTGGTG GCTTATGCCT GTAATCCTAG AACTTTGGGA 2640
GGCCTAGATG GGTGGATCAC TTGAGCTCAG GAGTTCAGAG CCAGCCCGGG CCTATGGCAA 2700
AACTCCGTCT CTACAAAAAA TAGAAAAAAT TAGCCAGGTG TGGTGGTGCA TGCCCTGTAGT 2760
CACAGTTTCA CCGCAGGCTG AGGTGGGAGG ATCACTTGAA CCCCAGAGGT CAAGACTGCA 2820
GTGAGCTGAG ATCAACACCA TGTACTCCAG CCTGGGTGAC AAAGTGAGAC TCTATCTCAA 2880
AAGAATAATT GGATCAATTT GTCAATTTCT ACAACACAA CAACAAAAAC CCTGTTGGG 2940
CACCTTGATT GAGATTGCAT TGAATTTATA TAAAACTGTT GGGAGAATTG ACATCTTAAT 3000
AATATTGATG CTCTGCGCCT ATAAACAAGG TCTGTCTTCC TAGGTATTAA TGTTTGTGCT 3060
TCTATTCTCT TTAATAATCT TTTGTAGTTT TCAGTGTACA GGTCTACCAT GTGAGCATTT 3120
CATAGTTTGG ATGCTAAATG GTATTTTAAA ATTTCAAAT CTAACCACT GTTGCTAGTA 3180
AATAGAAATA CAATTGATGT TGAACITGTA TCCTTCAGCC TTGCTAAACT GTGAGTTCTC 3240
ATGGTGTITT TGTAATTATC ATCAACAGTC ATGTGTTCTA TGAATAAAGA GTTTTACTCC 3300
TTC

Seq ID NO: 154 Protein sequence:
Protein Accession #: BAA11503.1

1 11 21 31 41 51
MFCEKAMELI RELHRAPEQG LPAFNEGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIIP 60
TIKFRHCSLL RNRRTVAVL YDRLLRIRAL RWEYGSVLPN ALRFHMAAEE MEWFNNYKRS 120
LATYMRSLGG DEGLDITQDM KPPKSLYIEV RCLDYGEFE VDDGTSVLLK KNSQHFLLPRW 180
KCEQLIRQGV LEHILS

Seq ID NO: 155 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 149-709

40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
GTTCCGCGCC AAAGCCGCGA GCGGAGGCCG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60
CGAAAGGAGT GAGGCGCCGA GAGCCAGAT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120
AAGGCCGCGG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180
CGAGCTGCAT CGCGCCCGCG AAGGGCAACT GCCTGCCCTC AACGAGGATG GACTCAGACA 240
AGTTCTGGAG GAGATGAAAG CTTTGTATGA ACAAACCCAG TCTGATGTGA ATGAAGCAAA 300
GTCAGGTGGA CGAAGTGATT TGATACCAAC TATCAAAATT CGACACTGTI CTCTGTTAAG 360
AAATCGAGCG TGCACTGTAG CATACTGTGA TGACCGCTTG CTTCCGATCA GAGCACTCAG 420
ATGGGAATAT GGTAGCGTCT TGCCAAATGC ATTACGATTT CACATGGCTG CTGAAGAAAT 480
GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTGAC TGGGAGGAGA 540
TGAAGGTTTG GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGCTGG 600
ATGCACTGTC GGCATCTCGG CTCAACCTGC AACCTCCACC TCCAGGTTTC ACCTCAACTG 660
CAACCTCCAC CTCCAGGTC CGGTGTCTAA AAGACTATGG AGAATTTGAA GTTGATGATG 720
GCACTTCAGT CCTATTAAAA AAAAAATAGCC AGCACTTTT ACCTCGATGG AAATGTGAGC 780
AGCTGATCAG ACAAGGAGTC CTGGAGCACA TCCTGTCTAT ACCATGCGCC GAGGCACCTC 840
CAGGCTTCAC TCAACTCATG GACTCCTCTG TACTCACTCT CTCCACCACCT CCCTTCACCT 900
CCCTCTTTGA TTTTAGAAGC TATAGACATT GTTTAAGATA ACTAAGAATA CTTGGCTAAG 960
AAGTATAATT TGCTAACTAT TAAGGACTTT CTTTTTTTAA TGTGTATAC TATTCTCTCT 1020
ACTCTTTTTC GGTTTTGGTT TTGTTTGTGA GAGACTGTCT CACTATGTTG CCAAGCTGG 1080
TCTCAAACCT CTGGCCTCAA GCAGTCTCTC CACCTTAGCT TCTCAAAGTG TTGAGATCAC 1140
AGGCGTGAGC CACTGCACCC GCGCCCTACT CCTTTTCTA ATAAGCTGTA TCTGTAATCA 1200
CAGCATCTCT ACAGTTGTTA CAGTGTGTTT TTTAAATGAA AGTAAACATG GTTACATTG 1260
AATCTCTTAA ATAAGCAGTC ACTTGGCTGG ACAGGAAGAA GSTAGATCCT GTGTGCTCTG 1320
TTTTCTGGTC ATGTGTATTG TACAAGCTAG AGAGCTGAAT TTCTGAGATA CACATTTTCA 1380
AATCACATGC AAGTGAAGAT GATGCTCTGT AGAAATTTTC AGTATATATA ATGTTTAATG 1440
ACATACTAAT TTATCATCTG GCTATTGGG AAGGAAGGAC ACACATGGAT TTGACACATT 1500
TCCACCATGG TGCGTGGTGT GGCTGTGGC TATGGGGTGA TCACCATGAT CACCACATTG 1560
GAAGGGGACA GTGAAATTGG GGCTAGAGAA GGAACCTTGT ACAGTTTTC CTGAGATTCA 1620
GATTGACTGA AAGTCACAT GAAGAGTTGA TTGCTTTTA ATGGTATGTT TTAACAGCT 1680
GACATTTTAA ATTTTGATGA AATCCAGTTT ATTCGTTTGT TCTTTTATGC TTTGGGTGTT 1740
GCATCCGAGA AATCTTTTCC CATCCCAAGA TCACAATTTT TTTTCTTTT TACTTCTAGA 1800
AGTGTATATA TTTTAAAGCT TATACCTTGG TCTATGACCC GTTTTCTTTT TTGTTTGTGTT 1860
TTGTTTTTTC GTTTGTTTCT TTGTTTGTAG ATGGAGTCTT GTTCTGTAC CCAGGCTGGG 1920
GTGCACTGCG GTGATCTTGG CTCACGTCAA TCTCTATCCC CTGGGTCAA GTGATTCTCT 1980
TGCTCTAGCC TCCCAAGTAG CTGGGATTAC AGGCACAGGC CGCCACGCTT GGCTAATTTT 2040
TGATTTTTTA GTAGAGACAG AGTTTTACCA TGTGGCCAG GCTGGTTTCA AACTCTGAC 2100
CTCAAGTGAC CCACCTTGGC CTCCCAAAGT TTTGGGATTA CAAGTGTGGG CCACCGCGGC 2160
CAGCCTATGA TCAATTTTGA ATGAATTTT TATATGGTGC AAGGTGTCAA TCCACCTTCA 2220
CTTTTCTTGT GGAATATAGA TATCCAGCTG TTTCACTACC ATTTTTTGA AGGACTGCCC 2280
TTTGCTCTAT CACCTTTGCA TTTTGTGTTA AAAGTAGTTG TCAATGTATA TGTGGGTTTA 2340
TTTCAGGACT CTGTTTGTCT CCATTGACCT GTTTTCTCT CCTGAATGCC AATACCATAT 2400
TTGTATGATG TGATATGTAAT TTTCTAATAA TTCTTGAAC AGATAGTATT AATGTGTCAAT 2460
ATTTTGTCTG GTTTTGTGAT TTTTGTAGA GATGGGGTTT CACCGTGTG GCCAGGCTGT 2520
GTTGAACCTC TGAGCTAAAG CAATACACTT GCCTCTGCTT CCCCATGTGC TGGGATTACA 2580
GGCGTGAGCC TTGGTGCTGG CCCAGTGATC CACATTTCTT TTTGAGATTT GTTTTGGCTA 2640
TGTTAAGTCC TTTGCTTTTG ATGTGAAAT TGGGAACAGG CAGGCTGTGG TGGCTTATGC 2700
CTGTAATCTC AGAATCTTGG GAGGCCTAGA TGGGTGGATC ACTTGAGCTC AGGAGTCCA 2760
GACCAGCCCG GGCCATATGC AAAACTCCGT CTCTACAAAA AATAGAAAAA ATTAGCCAGG 2820

5 TGTGGTGGTG CATGCTGTGA GTCACAGTTA CACGGCAGGC TGAGGTGGGA GGATCACTTG 2880
 AACCCAGAG GTCAGAGCTG CAGTGAGCTG AGATCACACC ACTGTACTCC AGCCTGGGTG 2940
 ACAAGTGAG ACTCTATCTC AAAAAGAAAT TAGGATCAAT TTGTCAATTT CTACAACAAC 3000
 AACAAACAAA ACCCTGTGTG GGCACCTTGA TTGAGATTGC ATTGAATTTA TATAAACTG 3060
 TTGGGAGAAAT TGACATCTTA ATAATATTGA GTCTCTGTGC CTATAAACAA GGTCTGTCTT 3120
 CCTAGGTATT AATGTTTTGT CTTCTATTTT TCTTAATAAT CTTTGTAGT TTTGAGTGA 3180
 CAGGTCTACC ATGTGAGCAT TTCATAGTTT TGATGCTAAA TGGTATTTTA AAATTTCAA 3240
 TTCTAACAC TTTGTCTAG TAAATAGAAA TACAATTGAT GTTGAACCTG TATCCTTCAG 3300
 CCTTGCTAAA CTGTGAGTTC TCATGGTGT TTTGTAAAT ACATCAACAG TCATGTGTTT 3360
 TATGAATAAA GAGTTTACT CCTTC

Seq ID NO: 156 Protein sequence:
Protein Accession #: Eos sequence

15 1 11 21 31 41 51
 MFCEKAMELI RELHRAPEQG LPAPNEDGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60
 TIKFRHCSLL RNRRTVAYL YDRLLRIRAL RWEYGSVLNP ALRFHMAAEE MEWFNNYKRS 120
 20 LATYMRSLGG DEGLDITQDM KPPKSLYIEA GCSGAISAQP ATSTSQVHLN CNLHLPGFVS 180
 KRLWRI

Seq ID NO: 157 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 148-621

25 1 11 21 31 41 51
 TTGGGCGCCA AAGCGCGGAG CGGAGGCGGA GGGAGAGGCC TGGCGCTGTA GGACTAGAAC 60
 GAAAGGAGTG AGGCGCCGAG AGCCCGAGATA CCATTTTGGC GTGAGAGCTG GTGGTTGGCA 120
 30 AGGCGCGCGG AGTGGGAAGC GTCCGCCATG TTCTGCGAAA AAGCCATGGA ACTGATCCGC 180
 GAGCTGCATC GCGCGCCGGA AGGGCAACTG CCTGCCTTCA ACGAGGATGG ACTCAGACAA 240
 GTTCTGGAGG AGATGAAAGC TTTGTATGAA CAAAACCACT CTGATGTGAA TGAAGCAAAG 300
 TCAGGTGGAC GAAGTGATTT GATACCAACT ATCAAAATTC GACACTGTTC TCTGTTAAGA 360
 AATCGACGCT GCACCTGTAGC ATACCTGTAT GACCGCTTGC TTGGATCAG AGCACTCAGA 420
 35 TGGGAATATG GTAGCGTCTT GCCAAATGCA TTACGATTTT ACATGGCTGC TGAAGAAGTC 480
 CGGTGTCTAA AAGACTATGG AGAATTTGAA GTTGATGATG GCACCTCAGT CCTATTAATA 540
 AAAAATAGCC AGCACTTTT ACCTCGATGG AAATGTGAGC AGCTGATCAG ACAAGGAGTC 600
 CTGGAGCACA TCCTGTCTAT ACCATGCGCC GAGGCACATC CAGGCTTCAC TCAACTCATG 660
 GACTCCTCTG TACTCACTCT CTCCACCACT CCCTTCACCT CCCTCTTTGA TTTTAGAAGC 720
 40 TATAGACATT GTTTAAGATA ACTAAGAATA CTGGCTAAG AAGTATAATT TGCTAATAT 780
 TAAGGACTTT CTTTTTTTAA TGTGTACAC TATCTTCTCT ACTTTTTTTT GGTTTTGGTT 840
 TTGTTTTGTA GAGACTGTCT CACTATGTTG CCCAAGCTGG TCTCAAACCT CTGGCCTCAA 900
 GCAGTCTCTC CACCTTAGCT TCTCAAAGTG TTGAGATCAC AGGCGTGAGC CACTGCACCC 960
 GGCCCTCCTC CCTTTTCTTA ATAAGCTGTA TCTGTAATCA CAGCATCTCT ACAGTTGTTA 1020
 45 CAGTGTGTTT TTTAAATGAA AGTAAACATG GTTACATTTG AATCTCTTAA ATAAGCAGTC 1080
 ACTTGGCTGG ACAGGAAGAA GGTAGATCCT GTGTGTCTTG TTTCTGGTC ATGTGTATTG 1140
 TACAAGCTAG AGAGCTGAAT TTCTGAGATA CACATTTTCA AATCACATGC AAGTGAAGAT 1200
 GATGGTCTGT AGAAATTTTC AGTATATATA ATGTTTAATG ACATACTAAT TTATCATCTG 1260
 GCTATTGTGG AAGGAAGGAC ACACATGGAT TTTGCACATT TCCACCATGG TGGCTGGTGT 1320
 50 GGCCTGTGGC TATGGGGTGA TCACCAGTAT CACCACITTG GAAGGGGACA GTGAAATTGG 1380
 GGTAGAGAAA GGAACITTTG ACAGTTTTC CTGAGATTCA GATTGACTGA AAAGTCACAT 1440
 GAAGAGTTGA TTGCTTTTAA ATGGTATGTT TTAACAGCT GACATTTTAA ATTTGTATGA 1500
 AATCCAGTTT ATTCGTTTGT TCTTTTATGC TTTGGGTGTT GCATCCGAGA AATCTTTTCC 1560
 CATCCCAAGA TCACAAATTT TTTTCTTTT TACTTCTAGA AGTGTATATA TTTTAAGCTT 1620
 55 TATACTTTGG TCTATGACCC GTTTTTTTTT TTGTTTTGTT TTGTTTTTTC GTTTGTTTCT 1680
 TTGTTTTGAG TGTGAGTCTT GTTCTGTGAC CCAGGCTGGG GTGCACTGGC GTGATCTTGG 1740
 CTCACTGCAA TCTCTATCCC CTGGGTTCAA GTGATCTCT TGTCTCAGCC TCCCAAGTAG 1800
 CTGGGATTAC AGGCACAGGC GGCACGCGCT GGCTAAATTT TGTATTTTAA GTAGAGACAG 1860
 AGTTTTACCA TGTGTGCCAG GCTGTTTCA AACTCCTGAC CTCAGTGAC CCACCTTGGC 1920
 60 CTCCCAAAGT TTTGGGATTA CAAGTGTGGG CCACCGCGGC CAGCCTATGA TCCATTTTGA 1980
 ATGAATTTTT TATATGTGTC AAGGTGTCAA TCCACCTTCA CTTTTTCTG GGAATATAGA 2040
 TATCCAGCTG TTTCACTACC ATTTTTTGAA AGGACTGCCC TTTGCTCTAT CACCTTTGCA 2100
 TTTTGTATAA AAGTAGTTG TCAATGTATA TGTGGGTTTA TTTGAGGACT CTGTTTGTGT 2160
 CCATGACCT GTTTTCTCT CCTGAATGCC AATACCATAT TTGTATGTAG TGTATGTAAT 2220
 65 TTTCTAATAA TTCTGAAAC AGATAGTATT AATGTGTCAT ATTTTGTCTG TTGTTGTAT 2280
 TTTTGTAGA GATGGGTTT CACCGTGTG GCCAGGCTGT GTTGAACCTC TGAGCTAAAG 2340
 CAATACACTT GCCTCGTCT CCCCATGTGC TGGGATTACA GGCCTGAGCC TTGGTGCTGG 2400
 CCCAGGTAC CACATTTCTT TTTGAGATTI GTTTGGCTA TGTAAAGTCC TTTGCTTTG 2460
 ATGTGAAAT TGGGAACAGG CAGGTTGTGG TGGCTTATGC CTGTAATCCT AGAACTTTGG 2520
 70 GAGGCTGAGA TGGGTGATC ACTTGAGCTC AGGAGTTCCA GACCAGCCC GGCCTATGGC 2580
 AAAACTCCGT CTCTACAAA AATAGAAAAA ATTAGCCAGG TGTGTGGTG CATGCCGTGA 2640
 GTCACAGTTA CACGGCAGGC TGAGGTGGGA GGATCACTTG AACCCAGAG GTCAAGACTG 2700
 CAGTGAGCTG AGATCACACC ACTGTACTCC AGCCTGGGTG ACAAAGTGAG ACTCTATCTC 2760
 AAAAAGAAAT TAGGATCAAT TTGTCAATTT CTACAACAAC AACACAAAA ACCCTGTTG 2820
 75 GGCACCTTGA TTGAGATTGC ATTGAATTTA TATAAACTG TTGGGAGAA TGCATCTTA 2880
 ATAATATGTA GTCTCTGGC CTATAAACAA GGTCTGTCTT CCTAGGTATT AATGTTTTGT 2940
 CTTCTATTTC TCTTAATAAT CTTTGTAGT TTTCAAGTGA CAGGTCTACC ATGTGAGCAT 3000
 TTCATAGTTT TGATGCTAAA TGGTATTTTA AAATTTCAA TTCTAACCA TTTGTCTAG 3060
 80 TAAATAGAAA TACAATTGAT GTTGAACCTG TATCCTTCAG CCTTGCTAAA CTGTGAGTTC 3120
 TCATGGTGT TTTGTAAAT ACATCAACAG TCATGTGTTT TATGAATAA GAGTTTTACT 3180
 CCTTC

Seq ID NO: 158 Protein sequence:
Protein Accession #: Eos sequence

85 1 11 21 31 41 51

MFCEKAMELI RELHRAPEQG LPAPNEDGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60
TIKFRHCSLL RNRRCCTVAYL YDRLLRIRAL RWEYGSVLPN ALRPFMAAEE VRCLRDYGEF 120
EVDGTSVLL KNSQHFPLR WKCEQLIRQG VLEHILS

5 Seq ID NO: 159 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 149-229

10 1 11 21 31 41 51
| | | | | |
GTTCCGCGCC AAAGCGCGGA GCGGAGGCC AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60
CGAAAGGAGT GAGGCGCCGA GAGCCACAGT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120
AAGSCGCGG GAGTGGGAAG CGTCCGCCAT GTTCTCGGAA AAAGCCATGG AACTGATCCG 180
CGAGCTGCAT CGCGCGCCCG AAGGCAACT GCGTGCCTTC AACAAATTAGC TGGGTGTGGT 240
GGCACACACC TGTAGTCCCA GCAACTTAGG AGGCTGAAGT GAGAGGATTG CATGGCTCCA 300
GGAAGTTGAA ACTGCAGTGA ACTGTGGTCA CGCTATTACA CTCCAGCCTG GGTGACAGAC 360
TGAATCCCTG TCTCAAAAGG GAAAAGGAGG ATGGACTCAG ACAAGTTCTG GAGGAGATGA 420
AAGCTTTGTA TGAACAAAAC CAGTCTGATG TGTCTCTGT TAAGAAATCG ACGTGCCTAC 480
GTAGCATACC TGTATGACCG CTTGCTTGG ATCAGAGCAC TCAGATGG

20 Seq ID NO: 160 Protein sequence:
Protein Accession #: Eos sequence

25 1 11 21 31 41 51
| | | | | |
ATGTTCTGCG AAAAAGCCAT GGAAGTATC CGCGAGCTGC ATCGCGCGCC CGAAGGGCAA 60
CTGCCTGCCT TCAACAATTA G

30 Seq ID NO: 161 DNA sequence
Nucleic Acid Accession #: U10694
Coding sequence: 1333-2280

35 1 11 21 31 41 51
| | | | | |
GGATCCGGCC GGATCTCAGG GAGGTGAGGA CTTTGTCTCT AGAGGGTGTG TGTGGACAAA 60
ACAGGGAGGC CCTGTGTTGG ACAGACACAG TGGTCCAGG ATTGGAGAGC AGTCCAGGTG 120
AGGAACCTAA GGGAGGATCG AGGGTACCTC CAGGCCAGAG AAACCTCTCAG ATCAAGAGAG 180
TTTGCCCTGC CCTACTGTCT ACCCCAGAGA GCCCGGGCAG GCGTGTCTGC TGAGGTCCCT 240
CCTTTATCCT GGGATCACTG GTGTGCGGGA GGGCTGGCCT TGGTCTGAGG GGGCTGCACT 300
CAAGTCAGCA GAGGAGGGGT CCCAGGCCCT GCCAGGAGTC CAGGTGACAG CTGAGGGGAC 360
CCCACTCACC AAACACAGAG GACCTAGCCC CACCCTGCCC CTGTGTCTAG CTGAGGGGAG 420
CCGCTGGGTG GATGACTGCT CCTCACTTCC TCTTCAGGTG TCTCCTGGAG ATAGGGCCTC 480
AGGTCAACAG AGGGAGGGTT CCAGACCCCTG CAGGCATCAA GATGAGGACC AGGCAGTATC 540
CTACCCCAAG GACACATGGA CCCCATTTGAA TTTAGACATC TCTTACTGTA CTTCCGAGGA 600
AACCCCTGGC AGGTGTGGGC AGATGTTGGT TGGGGCATGT CCTTCTGTT CATATCAGGG 660
ATGTGAGCTC CTGATCTGAG AGACTCTCAG GCAAGTAGAG GAGTAGAGTC CAGTCCCTGC 720
CAGGAGAAAG GTCAGGGGCC TGAGTGAGCG CAGAGGGGAC CATCCACCCC AAAAGTGTGT 780
AGAACTCAAG AGTGTCCAGC CGCCCTCTT GACAGCACTG AGGGACCGGG GCTCTGCCTG 840
CAGTCTGACG CCTAAGGGCC CTCTGATTCC TCTTCCAGGA GCTCCAGGAA GCAGGCAGGC 900
CTTGGTCTGA GACAGTGTCC TCAGGTGCGA GAGCAGAGGA GACCCAGGCA GTGTGAGCAG 960
TGAAGGTGAA GTGTTACACC TGAATGTGCA CCAAGGGCCC CACCTGCCCC AGCACACATG 1020
GGACCCATA GCACCTGGCC CCAITCCCCC TACTGTCACT CATAGAGCCT TGATCTCTGC 1080
AGGCTAGCTG CAGCTGTAGT AGCCCTCTCA CTTCTCCCT CAGGTTCTCG GGCACAGGTA 1140
ACCAAGAGGA CAGGAGCCCC AAGAGGCCCC AGAGCAGCAC TGAAGAGAC CTGTAAAGTCA 1200
GCCTTTGTTA GAACCTCCAA GGTTCGGTTC TCAGCTGAAG TCTCTCACAC ACTCCCTCTC 1260
TCCCAGGGCC TGTGGGTCTC CATGCCCCAG CTCTGCCCCA CGCTCCTGAC TGCTGCCCTG 1320
ACCAGAGTCA TCATGTCTCT CGAGCAGAGG AGTCCGCACT GCAAGCCTGA TGAAGACCTT 1380
GAAGCCCAAG GAGAGGACTT GGGCTGATG GGTGCACAGG AACCCACAGG CGAGGAGGAG 1440
GAGACTACCT CCTCTCTGA CAGCAAGGAG GAGGAGGTGT CTGCTGCTGG GTCACTAAGT 1500
CCTCCCCAGA GTCCCTCAGG AGGCGCTTCC TCCTCCATTT CGGTCTACTA CACTTTATGG 1560
AGCCAATTGG ATGAGGGCTC CAGCAGTCAA GAAGAGGAAG AGCCAAGCTC CTGGTGCAG 1620
CCAGCTCAGC TGGAGTTCAT GTTCCAAGAA GCACTGAAAT TGAAGGTGCG TGAGTTGGTT 1680
CATTTCTGCG TCCACAAATA TCGAGTCAAG GAGCCGCTCA CAAAGGCAGA AATGCTGGAG 1740
AGGCTCATCA AAAATTACAA GCGCTACTTT CCTGTGATCT TCGGCAAGC CTCGAGTTT 1800
ATGCAGGTGA TCTTTGGCAC TGATGTGAAG GAGGTGGAAC CGCGCGGCCA CTCCTACATC 1860
CTTGTCACTG CTCTTGGCCT CTGTCGAT AGCATGCTGG GTGATGTGCA TAGCATGCC 1920
AAGGCGCCCC TCCTGATCAT TGTCTGGGT GTGATCCTAA CCAAGACAA CTGCGCCCT 1980
GAAGAGTTA TCTGGGAAGC GTTGAGTGTG ATGGGGGTGT ATGTTGGGAA GGAGCACATG 2040
TTCTACGGGG AGCCAGGAA GCTGCTCACC CAAGATTGGG TGCAAGAAA CTACCTGGAG 2100
TACCGGCAGG TGCCCGGAG TGAATCTGCG CACTACGAGT TCCTGTGGGG TTCCAAGGCC 2160
CAOCTGAAA CCAGCTATGA GAAGGTGATA AATTATTGG TCATGCTCAA TGCAAGAGAG 2220
CCCATCTGCT ACCCATCCCT TTATGAAGAG GTTTTGGGAG AGGAGCAAGA GGGAGTCTGA 2280
GCACAGCCG CAGCCGGGGC CAAAGTTTGT GGGGTGAGG CCCCATCCAG CAGCTGCCCT 2340
GCCCATGTG ACATGAGGCC CATCTCTGCG TCTGTGTTG AAGAGAGCAA TCAGTGTCT 2400
CAGTGGCAGT GGGTGAAGT GAGCACACTG TATGTCTCT CTGGGTCTCT TGTCTATTGG 2460
GTGATTGGA GATTATCTCT TGCTCCCTTT TGAATGTGT CAAATGTCT TTTAATGGTC 2520
AGTTTAAATG ACTTACCAT CGAAGTTAAT GAATGACAGT AGTCACACAT ATTGCTGTTT 2580
ATGTTATTGA GAGTAAGAT TCTTGCTTT GAGTCACATG GGGAAATCCC TGTATTTTG 2640
TGAATTGGA CAAGATAACA TAGCAGAGGA ATTAATAATT TTTTGAAGC TTGAACITAG 2700
CAGCAAAATA GAGCTCAATA AGAATAGTG AAATGAAAT GTAGTTAAT CTGCTTAT 2760
ACCTCTTCT CTCTCTGTA AAATTAAC ATATACATGT ATACCTGGAT TTGCTTGGCT 2820
TCTTTGAGCA TGTAAAGAA ATAAATATG AAAGAATAAT TTTCTGTT CACTGGCTCA 2880
TTTTTCTTC AGACAGCAC TGAACATCTG TTATTCGAA CACCCTGGGT T

85 Seq ID NO: 162 Protein sequence:
Protein Accession #: AAA68877.1

1 11 21 31 41 51
 5 MSLEQRSPHC KPDELEAQQ EDLGLMGAQE PTGEEETTS SSDSKKEEVS AAGSSSPQPS 60
 PQGGASSSIS VYITLWSQFD EGSSSQESEE PSSSVDPALQ EFMFQELAKL KVAELVHPLL 120
 HKYRVKEPVT KAEMLESVIK NYKRYFPVIF GKASEFMQVI FGTDVKEVDP AGHSYILVTA 180
 LGLSCDSMLG DGHSMPKAAL LIIVLGVILT KDNCAPEEVI WEALSMVGIVY VGKHEMFYGE 240
 PRKLLTQDWV QENYLEYRQV PGSDPAHYEF LWGSKAHAET SYEKVINYLIV MLNAREPICY 300
 PSLYEVLGE EQEGV

Seq ID NO: 163 DNA sequence
 Nucleic Acid Accession #: AF292100
 Coding sequence: 30-809

1 11 21 31 41 51
 15 GGGGGGGGAG AGGCCCTGGAG GACACCAACA TGAACAAGTT GAAATCATCG CAGAAGGATA 60
 AAGTTTCGTCA GTTTATGATC TTCAACAAT CTAGTGAAAA AACAGCAGTA AGTTGTCTTT 120
 20 CTCAAAATGA CTGSAAGTTA GATGTTGCAA CAGATAATTT TTCCAAAAT CCTGAACCTT 180
 ATATACGAGA GAGTGTAATA GGATCATTGG ACAGGAAGAA GTTAGAACAG CTGTACAATA 240
 GATACAAAGA CCTCAAGAT GAGAATAAAA TTGGAATAGA TGGCATACAG CAGTCTCTGT 300
 ATGACCTGGC ACTCGATCCA GCCAGCATTG GTGTGTTGAT TATTGCGTGG AAGTTCAGAG 360
 CAGCAACACA GTGCCAGTTC TCCAACAGG AGTTCATGGA TGGCATGACA GAATTAGGAT 420
 25 GTGACAGCAT AGAACAACTA AAGGCCCAGA TACCCAAGAT GGAACAAGAA TTGAAAGAAC 480
 CAGGACGATC TAAGGATTTT TACCAGTTTA CTTTAAATTT TGCAAAGAAAT CCAGGACAAA 540
 AAGGATTAGA TCTAGAAATG GCCATTGCCT ACTGGAACCT AGTGCTTAAT GGAAGATTTA 600
 AATTCTTAGA CTTATGGAAT AAATTTTGTG TGGAAACATCA TAAACGATCA ATACCAAAAG 660
 ACACITGGAA TCTTCTTTTA GACTTCAGTA CGATGATTGC AGATGACATG TCTAATTATG 720
 30 ATGAAGAAGG AGCATGGCCT GTTCTTATTG ATGACTTTGT GGAATTTGCA CGCCCTCAAA 780
 TTGCTGGGAC AAAAGATACA ACAGTGTAGC ACTAAAGGAA CCTTTAGAA CTGTACATAGT 840
 CTGTACAATA AATACAACAG AAAATTCAC AGTCAATTC TGCTGGCTGG ACTGAACTGA 900
 AGATCAATCC TCACAATCA GACTGAGGCT TGAGACAAAA CTTTAAAGAT ACATCTTGGA 960
 CCATATCGTA TTTCATTCTT CTAATGGTGG TTTGGGCTTG TCTTCTAGTC TGGGCCGCTC 1020
 35 TAAACATTTA TAATTCACAC ATTGTGGATT TCATCTTATA TCTGTGGACC ATCCTAGTTT 1080
 ATTCCTCATC TAAGCTTAGA AGCTTTATGG TGATTATTTT GAGGTTTTC TCTGCGATA 1140
 AAGCACAATG CTGTCTTCAT CAGAAAACAG TTGGCATAAG AATTAACAT ATGAACATCA 1200
 CAAAACAATT TATAAAACT TCTTAAATAT ACCTTTGGG CTAGTTGCAA AGACTATGCT 1260
 40 AATAGCACTT CCAGTGAGAG TGATATATTT AAGTGTACTG GATCTGSAAT GGTGTTTGTG 1320
 TTTGGGGGGA ATTTTTTTTT TTCTCTGCA AATCACAAT GTTGTGATG TGAGTATCTG 1380
 ATGAAAAAAC AATGTCAAGAA TAACCGACAT GAAAATTTT TAGGATAACT TGGTGCTAC 1440
 CTGAAAAATG TATGTGTTT TAGACTCTTG ATTTCAAAAG GTTCCACAGA ACTAGTCTGC 1500
 GCTTACCTTA CCTAGTTTTA TATATAGCTG TCCTACAGGG AGCTTTTATT TAGAAAAATG 1560
 45 CTGCATAATG TTAGATTCTT CTCCTGTCTA CATTATGCAC TACATAATTG GACTTCATTA 1620
 TGCTTTTGAA ATGCTTATCT GCCTGTCACTA TAAGTTAAAC TATTTAAATTT GTTTTGAATG 1680
 TTTTGGATTG TAGACAATA CAATATTCTA AATTTAGGCA TGAGGGTTTT TTTGTTTTAT 1740
 TTTTACTTTT TTTTGTGCTA TGCACTATGG AACACAAATG AAATCTCTTT AATTTATAAG 1800
 AAGATAGTAG GAGTTAAATG TTGAAAATGG TTGTGATGAG CCACGAAAT CAATCTTTAT 1860
 50 AATATAGGTA CTGCTCTTTC AGACAAACAG TCCATTTTTA ATGACTTCTT ATTTTGTGTA 1920
 AATTACTTTA ACTGCTAATC ACTGTGGTTG CCAAAATATT ACTTCAGAAG CAAAGATTTT 1980
 CAAACAAGCA AATGTCAAGC AAAAATACAG TCTGGCTTCT AGTCTATTTA CTGTTTGTGT 2040
 TCACTCAGAT TAGCTCAGTT TTCTCATCAA AGCAGAAATG TATCTTGGCT GTGTGTGTGT 2100
 GGTGTGTGT GTGTGTGTGT GTATGTGTGT ATATATATAT ATATATATAT ATATATATTT 2160
 55 TTTTTTTTTT TTTTTTTTAA ATTACAAAAG CCATGAGCTG CTTTATAGCT GAAAATGGTC 2220
 ATTTCCCTGT TCACCTACTG ACATGTGAAG AAGGGTTTCT TGCTTTCTTA AACATTTCCG 2280
 TAAGGCAGGC TAGAAATGTA ATACTTCAA TGTTTGATGA TTATGCTCTT TTGATAGGAA 2340
 TAGATTCTGC TTGGGATATA TATCCAGGCA CTCTCTAAGG TCTAGGGTTG ATATTAAACA 2400
 AGGAATGTAC TTAGAATAGC AGTACATTTT ATGCAAAATAT GSAATTAAT TTAAGAAACA 2460
 60 ATGACATATC AAAACTGCTT TTTACATGAT TTTGAAATAG ACTAGAAAGC TTTCCCTATA 2520
 GACATATTAA TATTCACATC ATAACCTTAA TTCAAGAAATG CAGTTTAC AAAGAAAAAA 2580
 TTTGAAAATT TCTATTACGG CTACTGGAAT TGGTTATTAA AAGAAAAAGG AAAAGAAAGA 2640
 ATCTGTCTGC TTTCAATATT TCCTGATTTT TTTGTAATAA TAAAGAGGAA CTTCAATTAT 2700
 65 GAAAAATTTT TAAAGATAT ATATATCTAT ATATCTATAT ATATGACTG TTTTGTTCCT 2760
 TGCTTTGAAG ATTTTGAGTT ATGGTTATTG GTTTCAGATT GATTAATTC CATATGCTGT 2820
 GTTTTCTTTA AAGTCATAT GGGTTCGTGG CCTAATGCCT TGGAATTTAC ATATTTTCT 2880
 TTTTAAATGC AAAACCTTTT CAACAAAATA GTGTTTGTC TCAAGTTTGT ACTAAACATT 2940
 70 TATAATTACT GTGTAATTAT AAACAAAAT ACATAAAGCT TTGAATATAA TTATGTAGCA 3000
 TAAAAGTTAA GGTGTGTCAC TATGATGGCA TCTTAGAATT AAACAAAAT TTTACTAGGG 3060
 CTGAAAAGAG AAGACTGATT TAATGTGGTG TGATTATTCT GAAGATAAAT GTCTGGCTAC 3120
 AGGGAATATT TTGACTAAA AAATGATTAC ACATATGGCT GTGTGTGTTT GAGTCTGTGT 3180
 CTGTGAGAGA GCCAGAGAGA GTGAGAGAGA TTGACAGAGA AAGGGAGAGA CACACACAG 3240
 75 CCCCCTGAAT TGCTTAACT CCTAAGTGT TCACTCCTCA TTCCGGTAAA CCCCCATGC 3300
 TGATCTTTG TTTTAAACTG AACCATAGGT ACAGTTTCCCT TTTTGCCAAA TGTCAAAACA 3360
 GGTACAAATT TTAAATGTGA ATGCTTTTFA AATAGAAAAA TGTATAAAAT TAGAAGTGCC 3420
 CACATATAAA AAATACTTGA GATGAAGATT ATCTTTAGTG AATATCATCT GCATATCTCT 3480
 80 GTAAAGTTCAA TTGTGTTCT TACAGTCCCT GTCATATTAC CAACAGAGGC AATAAAAGCT 3540
 GCAGTGAAT TG

Seq ID NO: 164 Protein sequence:
 Protein Accession #: AAG00606

1 11 21 31 41 51
 80 MNKLKSSQKD KVRQFMFTQ SSEKTAVSCL SQNDWKLDVA TDNFFQNPFL YIRESVKGSL 60
 DRKLEQLYN RYKDPQDENK IGIDGIQFQC DDLALDPASI SVLILIAWFR AATQCEFSKQ 120
 85 EFMGMTELG CDSIEQLKAO IPKMEQELKE PGRFKDFYQF TFNFARNPGO KGLDLEMAIA 180
 YWNLVLNRF KPLDLWNKFL LEHKRSIPK DTWNLLDFPS TMIADDSMNY DEEGAWPVL 240
 DDFVEFARPO IAGTKSTTV

Seq ID NO: 165 DNA sequence
Nucleic Acid Accession #: AF256215
Coding sequence: 220-2028

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1	CTCCAGTCCG	11	CATGCTCAGT	21	AGCTGCTGCC	31	GGCCGGGCTG	41	CGGGGCGGCG	51	TCCGCTGCCG	60
	GCCTACGGCG		TGCGGTGGCG		GCCGCGCGCG		CACCGCGCAG		GGCCCGCCAG		TCCCGCTTTC	120
	CCTGCTCCAG		AGCGCGCGCG		TGGGCGCGGG		CAGGCGCGGC		CGGGGGCTCC		TCCATGCTGC	180
	CAGCGCGCGG		GCTCGCGAGC		CGACCAAGTG		GCTCCTGGGA		TGGCGGCGGA		AGAGGAGSCT	240
	GCGGCGGGAG		GTAAGTGT		GAGAGAGGAG		AACCAAGTGA		TTGCTCTGT		GGTTTCAGC	300
	CGCGTGAGTC		CAGGAGCAAG		ACCAACAGCT		ATGGGCTCTT		TCAGCTCACA		CATGACAGAG	360
	TTTCCACGAA		AACGCAAGG		AAGTGATTCA		GACCCATCCC		AAGTGGAAGA		TGGTGAACAC	420
	CAAGTTAAAA		TGAAGGCCCT		CAGAGAAGCT		CATAGCCAAA		CTGAAAAGCG		GAGGAGAGAT	480
	AAAATGAATA		ACCTGATTGA		AGAATGTCT		GCAATGATCC		CTCAGTGCAA		CCCCATGGCG	540
	CGTAAACTGG		ACAACTTAC		AGTTTAAAGA		ATGGCTGTTT		AACACTTGAG		ATCTTTAAAA	600
	GGCTTGACAA		ATTCTTATGT		GGGAAGTAAT		TATAGACCAT		CATTCTTCA		GGATAATGAG	660
	CTCAGACAT		TAATCCTTAA		GACTGCAGAA		GGCTTCTTAT		TTGTGGTTGG		ATGTGAAGA	720
	GGAAAAATTC		TCTTCGTTTC		TAAGTCAGTC		TCCAAAATAC		TTAATTATGA		TCAGGCTAGT	780
	TTGACTGGAC		AAAGCTTAT		TGACTTCTTA		CATCCAAAAG		ATGTTGCCAA		AGTAAAGGAA	840
	CAACTTTCTT		CTTTTGATAT		TTCAACCAAGA		GAAAAGCTAA		TAGATGCCAA		AACCTGGTTG	900
	CAAGTTTACA		GTAATCTCCA		CGCTGGAAAG		ACACGTGTGT		ATTCTGGCTC		AAGACGATCT	960
	TTTTTCTGTC		GGATAAAGAG		TTGTAAATAT		TCTGTCAAAG		AAGAGCATGG		ATGCTTACCC	1020
	AACTCAAAGA		AGAAAGAGCA		CAGAAAATTC		TATACTATCC		ATTGCACTGG		TTACTTGAGA	1080
	AGCTGGCCCT		CAATATTGT		TGGAATGGAA		GAAGAAAGGA		ACAGTAAGAA		AGACAACAGT	1140
	AAATTTACCT		GCCTTTGGGC		CATTGGAAGA		TTACAGCCAT		ATATTGTTCC		ACAGAACAGT	1200
	GGAGAGATTA		ATGTGAAGCC		AAGTGAATTT		ATAACCCGCT		TTGCACTGAA		TGGAATTTT	1260
	GTCTATGTAG		TAATCAAGGC		AACAGCGATT		TTAGGATATC		TGCTCAGGA		ACTTTTGGGA	1320
	ACTTCTGTT		ATGAATATTT		TCATCAAGAT		GACCACAATA		ATTTGACTGA		CAAGCACAAG	1380
	GCAGTTCTAC		AGAGTAAGGA		GAAAATACTT		ACAGATTCCT		ACAAATTCAG		AGCAAAAGAT	1440
	GGCTCTTTTG		TAACTTTTAA		AAGCCAATGG		TTTAGTTTCA		CAATCTCTTG		GACAAAAGAA	1500
	CTGGAATATA		TTGATCTCTG		CAACACTTTA		GTTTGGGAGC		ATAGTGAGCC		TGGAGAAGCA	1560
	TCATTTTAC		CTGTAGCTCT		TCAATCATCA		GAAGAATCCT		CTAGACAGTC		CTGTATGAGT	1620
	GTACCTGGAA		TGTCTACTGG		AACAGTACTT		GGTGTCTGTA		GTATTGGAAC		AGATATTGCA	1680
	AATGAAATTC		TGGATTTTACA		GAGGTTACAG		TCTTCTTCAT		ACCTTGATGA		TTGAGTCCA	1740
	ACAGGTTTAA		TGAAGATATC		TCATACTGTA		AACCTGAGGA		GTATGTCAAA		TAAGGAGTTG	1800
	TTTCCACCAA		GTCTCTCTGA		AATGGGGGAG		CTAGAGGCTA		CCAGGCCAAA		CCAGAGTACT	1860
	GTGCTGTGTC		ACAGCCATGA		GCCACTCCTC		AGTGATGGTG		CACAGTTGGA		TTTCAGTGCC	1920
	CTATGTGACA		ATGATGACAC		AGCCATGGCT		GCATTTATGA		ATTACTTAGA		AGCAGAGGGG	1980
	GGCCTGGGAG		ACCTGGGGGA		CTTCAGTGAC		ATCCAGTGGA		CCCTCTAGCC		TTTGATTTTT	2040
	AACTCCAAAA		ATAGAAAAAC		TTTTAAAGCA		TTATTACGA		AAAAACTGTC		TCAACTATTC	2100
	TTAAGTACTG		TATTGATATT		GTTTGTATCT		TTTATTAAAT		TTCTACCACT		TTTTATAGAT	2160
	TGTCATCTTC		CTGTACAGG		GATGTGGGGA		AATACGTTTT		CCTCCCAAGA		GAACCAAGTT	2220
	TATTATAGAC		TCTTTTATCT		AGTGAAATGG		CTTATAATCC		ACTAGTTGCC		ATATTTTTCG	2280
	TAAATATATT		CTAACCAAGA		ATACTACTTA		CATATTGTTT		TGGCTTTGTT		TTATTTTGA	2340
	TGCACTTTTG		TTTAGTTGAG		GTAATGTAAT		ATATTGATGT		TTTCTTTTGT		GTCTAAGATT	2400
	GATTATATAT		AGTAGGTTTG		TATAATTGCG		AACATTTTCC		ATGCTTGGCG		AATTTCCCTA	2460
	ATTGAGGATA		GGGCTTACAC		ACTTTAAGAA		AACAGTGAGT		ACTTGAAAT		TAAAGGGAC	2520
	AGTGCAATT		ATAGTCATAA		TCACATTGAA		TACTGTATTT		GATCTTTGGA		GACTTAGGCA	2580
	AGCACAGAGC		TGGGATATTT		ATGCTCAGTT		GAGCACTTTA		AGATGAATTT		TAAGTGAGAT	2640
	GATTTCTTTG		TTAAAGTCAA		GAAAGTCAAA		AGAGTTTCAG		CTTCTCTTAC		AGAAAAGGAA	2700
	GGATCTTGGG		CCCTAGATCT		TGGGGATTAA		CCTCTGCATA		TAAGATTATC		TCTTAATAGG	2760
	CCAGACGTGG		TGCTCACGGC		TGTAATCCCA		GTACTTTGGG		AGGCTGAGAC		GGGCAGATCA	2820
	CTTGAGGTCA		GGAGTTCAGG		ACCAGCCTGG		CCAATATGGT		GAAACCCCGT		TTCTACTAAA	2880
	AATACAAAAA		AAATTAACCA		GGCACTCACT		CTTGAGGTAA		CTAACCAACT		CCACGATATA	2940
	TGACAGTCTA		TTCTATGAGC		CAAGGGCCTC		ATGACCTAAT		GGCACACACC		TGTAATCCCA	3000
	ACTGCTTGGG		AGGCTGAGGC		GAGAGGATTG		CTTGAACCTG		GGAGGCAGAG		GTTCAGTGA	3060
	GCCGAGATCG		CACCACTGCA		CTCCAGTCTG		GGCAACAGAG		TGAGACTTCA		TCTCAAAAAA	3120
	AGTAAAAAAA		AAGATTTAAT		ATAATCACTG		AAGATCTCTA		TTATAGATAG		ATTAGGTTTT	3180
	TGACATTTGA		AACATCTTAA		GGGATAGATT		TGTCTCTAAG		GAAAAAAGTA		GGCCCGGGCA	3240
	GATTAAATGT		CTTGTGTAAA		GTCAACATTT		AAATTCAGTC		ACACATTAAA		TTCTATAGAT	3300
	TTTAAATGTT		TAAATGATAT		AAACCAAGTT		CITTATACAC		ATTTGGGAAA		ACATTGGTCT	3360
	CACAGATTAA		ATGATTAACT		AACAGACCCA		GGAACTAGTT		GTAGCTTTCT		AAGTAATTAG	3420
	GCAATTACAG		TTATTGCTCT		TAAACCAAGG		TAATAAAACA		AAATGACAAG		TACATGTTTA	3480
	AAATTATGAG		GCAATGAGAA		ATAATTTAAA		AACCAATTTT		CTAGTTATAA		TTTAAATTTT	3540
	GGAGAGCATT		TTTAAAGTAA		ATTAATCCAG		AGGTGGCTCA		AATTGAGTAT		AAGAATTAAAG	3600
	ATTATTTAAA		ATACTGATG		TCTACCTTCT		CGGGGATCAT		ACTTTATAAC		ACTTCTGCT	3660
	TCAGTAGCTC		TTATAGCTT		GCCAAATATG		CTCCCATATT		TTCTCTCTCG		TGCCCTCGCA	3720
	ATGAAAGTCA		GATAGGCTGG		GAACTCATGG		GGCAGCCCTC		AGACTTCAAT		GTGGGCTTCA	3780
	AATCCAGTTT		CCTGTTCTAT		ATGGTGCTAC		ATCTTTCCAG		AAAAATTTCC		TCAGAGCCCC	3840
	TCGCCAAAAA		AAAGCAATTAT		TTTGACCCCTG		CATGCTATTT		CTTTAGCTGT		AGGTGATAGA	3900
	TTAGAACTTC		TGTCAGACAT		GTTAATGACA		AACATACCAA		CAGACAATAA		CCAAAGCAAA	3960
	TGTTTCTCTT		AAGTGTGAAA		TGTGCAGGGG		CTCGTGGGCA		AGGATGTATT		GGCACACTGT	4020
	CCTCTTGAAC		TGATAGTGT		CCAGCAATGT		TGGAGGTTGG		CACCATTCCT		GGTCCGACAC	4080
	TTAGGAGCTT		GAGAGACATC		AGGTTTAGAA		TGAGCCAAAG		AAATCCTACA		AGATGGGGAG	4140
	AAATTGGTGT		CAGCAGCCTA		AGTGTATAG		TTAAGTCTAA		AGAAGTATGA		AAGATCCOCT	4200
	GTGTTCTCTA		AATTGAGCAG		AGGGGCTGCT		CTACCAATAT		CACTTTITAG		GGGACTGAAC	4260
	CATTGCAAGT		TAGACTTTGG		TTCCAAAGAG		TCTGCCTAAG		CCAGGGGTGG		CAGGATAGGC	4320
	CATCATAGCT		GGATGGCTCT		AAAAGCAGAT		GGGGGCGAGC		TTGCCCTCGT		GATGCCAGGA	4380
	TTTGAGAGCG		AGAGTTTCTA		GAGGGAGACC		AGTGCTGCTT		CTCACAGTGG		CAGTTTITTC	4440
	TCCTTGGCAAG		AGGAGGGGCT		GTTCAATTCC		ATAGACCACT		GGGCAGATAG		CCAGTTGAAT	4500
	ACTCTGTGCA		TGGTTTGATC		CTTTAATAGT		TGCTCTAAT		ATTTTCTGT		AGATCCITTT	4560
	GTCTCTGACT		CAAAATCTAA		TCCATGCATT		GTATGATACC		GTAGCTCTCC		TAAGGTTTGT	4620
	GTTCCTCTCA		AAATGTTTTTA		GTTTCTTCTA		ACTAAATTTG		ATTTTGTGCT		TTAGAAGTGA	4680
	CATATTTTAA		TGGTATACAC		TATGTTCCCT		TTTCTACTG		CGAGTCAATT		TTTGAATTT	4740
	TCGTGAGAAA		GAATATATCT		ACAAATTGCA		CGAAAGTATC		ATAAAAACAG		TACTCTAGAG	4800

	CAGCGCTGTC	CAATAGAAAT	ATAATCTGAG	CCACATGTAT	AATTTTATTT	TCTTCTAGCC	4860
	ACATTAAGA	AGTAAAAAGA	TACAAAGTAGA	ACTAATTTTA	ATGTTTAAAT	TCAGTATATC	4920
	CAAAATATCA	TTTGAACATG	TAAATTAATAT	AAAATTATTA	ATGTGATATT	TTACATTCTT	4980
5	TTGGAATAC	TAGTCTTCAA	AATCTGGTAT	GTATCTTACA	TTGATAGCAC	ATCTCACTTT	5040
	GTACTAGCCA	CATTGCAAGT	GCTCAGTAGC	CACATGTGGC	TAGTGGCTAC	TGCACTGGAC	5100
	AGCACAGTTC	TAGGTTCCAC	CCTAACACCC	AAGTCCTGTG	GATTAGAATC	CCAGAAATCAG	5160
	AGCTGGAAAT	AAACATAGAG	ATCAAAACCTC	CTTTTAAAAA	TGAGGACGCT	GAGGCACAGA	5220
	GTTTAAATGG	CTTGATGAGC	GTACATACAGC	TAAATTGAGC	CTCAACAGGG	TCTTCTGATT	5280
10	CCAGGCACCT	TTCCCACTCC	ACTACATTAC	TGTAGTGGTA	ATCTTAGGG	TTAAAAAAG	5340
	TGTAGAGTAG	GCCGGGCGCA	GTGGCTCATG	CCTGTAATCC	CAGCACTTTG	GGAGGCGGAA	5400
	GTGGGCGGAT	CACGAGGTCA	GGAGATCGAG	ACCATCCTGG	CCAACATGGT	GAAACCCCGT	5460
	CTCTACTGAA	AATACAAAGC	AAAATTAGCC	AGGTGTGGTG	GCGGGCGCCT	GTGGTCCCGAG	5520
	CTGCTCTGGA	GGCTGAGGCA	GAATGGCGTG	AACTCCAGGAG	GCAGAGATGG	CAGTGAGCCA	5580
15	AGATGGCGCC	ACTGCACCCC	AGCCTGGGCG	ACAGAGCGAG	ACTCCATCTC	AAAAAAGAA	5640
	AAAAAAGAA	AAGAAAAGAA	AAGAAAAGTC	TAGAGAACAT	TATATTAAGT	GGTTATTATT	5700
	GAAGTAGACC	AAAGTTTATA	CCATAAGGAT	ATTTTCTCCT	AAATACCATG	TTTGAAGAAC	5760
	AAATTATTAT	TGATCCTTGA	ATCTGTAAGA	TCAAATAACA	AGTCTCTATC	CATGTTACCA	5820
	AAITTAACCT	TTTGAATAA	ATAAACTTTA	AAATATCAGA	TGTGTATTA	CAGGATGATA	5880
20	CTTGAATCA	AGTGAATGA	GTTATATGGT	CATCACTAAA	TTTAGAAATC	TATTGTGAAA	5940
	CAAGACAAA	CAGGAAGTAG	ACTTTTAGTA	AATAAATGGA	ATTTAAAGAA	ATTTAAAGAA	6000
	AAGTGTATAT	TTACAGTGGT	ACGACAGAAA	AGGATGTCTT	TGTTGTGATA	GTCTTTGAGG	6060
	GATCTCGTAT	AAATCTGGGG	CACAGGTACA	AGAAATAGCC	AATATTAGT	TCCAGAGCCA	6120
	TGTTTAGTAG	TGTCAGATT	CAGATCATGC	TGCCAAGAGG	TATCTCCCCC	TCAAGTGGGT	6180
25	CATCACTGAG	CCCTGGAAIT	GGAGACTCAT	ACTTGCCGAG	CACAAATGTA	CGGGCAGACA	6240
	GGCCGACATC	TATAGATTAGC	TAGAAGCCAT	AAAGAAAAGC	TGCTAAGTGG	CCAATAGGTG	6300
	CCACTTTTCT	GTCTTTGTAA	TGCTTTCAIT	AGCAGATCTT	TTTTTCCCAA	GCTCCATGGG	6360
	GCCTATGAGA	GGCATTATAT	ATTTTGTGTC	CTACAATAAG	TCAGCCTGTC	TGGTGTGAGT	6420
	TGTTTATAGA	GAATGCTTT	CCAAGGGAGG	TCTAGGAAGA	TCTGACACA	TAGAACTTT	6480
30	GGCTTAGAGA	GCTTTCCAGT	TGTAGTGCCA	ATAAAACTG	ACCTGGAAG	AAAACCTGCC	6540
	CAGCAGGAAA	CATGCTTTCT	GAACTCACTT	GAGAGTGTAT	GGTGTATGTC	ACTTCTCATA	6600
	TATCTTAGAG	TTTAGATTG	TCTTTTATAC	AATTTTATGC	TCTTTTCCAG	TTCACTTGTG	6660
	CTGCTCTGTA	TATGGTATT	TTTAAATTTT	TGTGGTAAAT	AATGAAAAGA	GTGAAATTAT	6720
35	ATTTTATAAT	TACTCATTTG	TAGTTTTTTT	TTTTAATTTA	ATAAACTTCC	TCCAAAAGT	6780
	GCTCCCTTAA	AA					

Seq ID NO: 166 Protein sequence:
Protein Accession #: AAG34652

40	1	11	21	31	41	51	
	MAAEERAAAG	GKVLREENQC	IAPVVSSRVB	PGTRPTAMGS	FSSHMTFEPF	KRKGSDSDPS	60
	QVEDGEHQVK	MKAFREAHQS	TEKRRRDKMY	NLIEELSAMI	PQCNPARKL	DKLTVLRMAV	120
45	QHLRLSLKGLT	NSYVGSNYRP	SFLQDNELRH	LILKTAEGFL	FVVGCEGRKI	LFVSKSVSKI	180
	LNVDQASLTG	QSLDFDLHPK	DVAKVKEQLS	SFDISPREKL	IDAKTGLQVH	SNLHAGTRV	240
	YSGSRSPFC	RIKSKCISVK	EEHGCLPNSK	KKEHRKFYTI	HCTGYLRSWP	PNIVGMEER	300
	NSKKDINSNPT	CLVAIGRLQP	VIVPQNSGEI	NVKPTEFTR	FAVNGKFVYV	DQRATAILGY	360
	LPQELLGTSC	YEYFHQDDHN	NLTDKHKAVL	QSKEKILTDS	YKFRADGGSF	VTLKSQWFSP	420
50	TNPWKELSY	IVSVNTLVLG	HSEPGESFL	PCSSQSSEES	SRQSCMSVPG	MSTGTVLGAG	480
	SIGTDIANEI	LDLRLQLQSS	YLDSSPTGL	MKDTHTVNCR	SMSNKELFPF	SPSEMGELEA	540
	TRQNSTVAV	HSHEPILLSG	AQLDFDALCD	NDDTAMAAFM	NYLEAEGGLG	DPGDFSDIQW	600
	TL						

Seq ID NO: 167 DNA sequence
Nucleic Acid Accession #: NM_014400
Coding sequence: 86-1126

60	1	11	21	31	41	51	
	GGTACTCAT	CCTGGGCTCA	GSTAAGAGGG	CCCAGGCTCG	GAGGCGGCAC	ACCCAGGGGG	60
	GAGCCCAAGG	GAGCAGGAGC	GAGCCATGGA	CCCGCCAGG	AAAGCAGGTG	CCCAGGCCAT	120
	GATCTGGACT	GCAGGCTGGC	TGCTGCTGCT	GCTGCTTCGC	GGAGGAGCGC	AGGCCCTGGA	180
	GTGCTACAGC	TGCGTGCGAG	AAGCAGATGA	CGGATGCTCC	CCGAACAAGA	TGAAGACAGT	240
65	GAAGTGGCGG	CCGGCGGTGG	ACGTCTGCAC	CGAGGCCGTG	GGGGCGGTGG	AGACCATCCA	300
	CGGACAAATG	TGCGTGCGAG	TGCGGGGTGG	CGGTTGCGGA	CTCCCCGCGC	AGAATGACCG	360
	CGGCGTGGAT	CTTCAAGGGC	TTCTGGCGTT	CATCCAGCTG	CAGCAATGCG	CTCAGATCGC	420
	CTGCAACGCC	AAGCTCAACC	TCACCTCGCG	GGCGCTCGAC	CGGCGAGGTA	ATGAGAGTGC	480
	ATACCCGCCCC	AACGGCGTGG	AGTGCTACAG	CTGTGTGGGC	CTGAGCCGGG	AGGCGTGCCA	540
70	GGGTACATCG	CCGCGGTGCG	TGAGCTGCTA	CAACGCCAGC	GATCATGTCT	ACAAGGGCTG	600
	CTTCGACGGC	AACGTCACTT	TGACGGCAGC	TAATGTGACT	GTGTCTCTTC	CTGTCCGGGG	660
	CTGTGTCAG	GATGAATCTT	GCACTCGGGA	TGGAGTAAAC	GGCCAGGGGT	TCACTCTCAG	720
	TGGCTCCTGT	TGCCAGGGGT	CCCGCTGTAA	CTCTGACCTC	CGCAACAAGA	CCTACTTCTC	780
	CCCTCGAATC	CCACCCCTTG	TCCGGCTGCC	CCCTCCAGAG	CCCAGGACTG	TGGCTCTAAC	840
75	CACATCTGTC	ACCACTTCTA	CCTCGGCCCC	AGTGAGACCC	ACATCCACCA	CCAAACCCAT	900
	GCCAGCGCCA	ACCACTTCTA	CTCGAGACCA	GGGAGTAGAA	CACGAGGCCCT	CCCGGGATGA	960
	GGAGCCGAGG	TTGACTGGAG	GCGCGCTGGG	CCAACAGGAC	CGCAGCAATT	CAGGGCAGTA	1020
	TCTTGCAAAA	GGGGGGCCCC	AGCAGCCCCA	TAATAAAGGC	TGTGTGGCTC	CCACAGCTGG	1080
	ATTGGCAGCG	CTTCTGTGTC	CGTGGCTGCG	TGGTGTCTTA	CTGTGAGCTT	CTCCACCTGG	1140
80	AAATTTCCCT	CTCACTTACT	TCTCTGGCCC	TGGGTACCCC	TCTTCTCATC	ACTTCTCTGT	1200
	CCCAACCATG	CTCACTTACT	TCTCTGGCCC	TGGGTACCCC	TCTTCTCATC	ACTTCTCTGT	1260
	CTTCTGCTCG	GCTGGTTTGC	GGCTTTGGGA	AATAAAATAC	CGTTGTATAT	ATTCTGGCAG	1320
	GGGTGTTCTA	GCTTTTGTAG	GACAGCTCCT	GTATCCTTCT	CATCCTTGTG	TCTCGGCTTG	1380
	TCTCTTGTGT	ATGTTAGGAG	AGAGTGAGAG	AAGTCAGCTG	TCAAGGGGAA	GGTGAGAGAG	1440
85	AGGATGCTAA	GCTTCTTACT	CACTTTCTCC	TAGCCAGCCT	GGACTTTGGA	GCGTGGGGTG	1500
	GGTGGGACAA	TGGCTCCCCA	CTCTAAGCAC	TGCTTCCCCT	ACTCCCCGCA	TCTTTGGGGA	1560
	ATCGGTTCCC	CATATGTCTT	CCTTACTAGA	CTGTGAGCTC	CTCAGGGGCA	GGGACCGTGC	1620
	CTTATGTCTG	TGTGTGATCA	GTTTCTGGCA	CATAAATGCC	TCAATAAAGA	TTTAATTACT	1680

TTGTATAGTG AAAAAAAA

Seq ID NO: 168 Protein sequence:
Protein Accession #: NP_055215

5
10
15

1	11	21	31	41	51	
MDPARKAGA	Q	AMIWAGWLL	LLLLRGGQA	LECYSCVQKA	DDGCSFNMK	TVKCAPGV
CTEAVGAVET	I	HGQFSLAVX	GCGSGLPGKN	DRGLDLHGLL	AFIQLQQCAQ	DRCNALNLT
SRALDPAGNE	S	AYPPNGVEC	YSCVGLSREA	CQGTSPPVVS	CYNASDHVYK	GCFDGNVTLT
AANVTYSLPV	R	GCVDQDEPCT	RDGVITGPGFT	LSGSCCQGSR	CNSDLRNKTY	FSFRIPPLVR
LPPPEPTTVA	S	TVVTSTTS	APVRPTSTTK	PMPAPTSTTP	RQGVHEEASR	DEEPRLTGGA
AGHQDRSNSG	Q	YPKAGGPQQ	PHNKGCVAPT	AGLAALLLAV	AAGVLL	

Seq ID NO: 169 DNA sequence
Nucleic Acid Accession #: NM_006875
Coding sequence: 186-1190

20
25
30
35
40
45
50
55

1	11	21	31	41	51	
GAATTGCGCA	CGAGCGCGCG	GCGAATCTCA	ACGCTGCGCC	GTCTGCGGGC	GCTTCCGGGC	60
CACCACTTTC	TCTGCTTTC	ACCCCTGGCGC	CCCCCAGCCC	TGGCTCCCCA	GCTGCGCTGC	120
CCCGGGCGTC	CACGCCCTGC	GGGCTTAGCG	GGTTCAGTGG	GCTCAATCTG	CGCAGCGCCA	180
CCTCCATGTT	GACCAAGCCT	CTACAGGGGC	CTCCCGCGCC	CCCCGGGACC	CCCACGCCGC	240
CGCCAGGAGG	CAAGGATCGG	GAAGCGTTGG	AGGCGGAGTA	TGCACTCGGC	CCCCTCTGG	300
GTAAGGGGGG	CTTTGGCACC	GTCTTCGCAG	GACACCGCCT	CACAGATCGA	CTCCAGGTGG	360
CCATCAAAAT	GATTCCCGCG	AATCGTGTGC	TGGGCTGGTC	CCCCTGTGCA	GACTCAGTCA	420
CATGCCCACT	CGAAGTCGCA	CTGCTATGGA	AAGTGGGTGC	AGGTGGTGGG	CACCTGGCG	480
TGATCCGCTC	GCTTGACTGG	TTTGAGACAC	AGGAAGGCTT	CATGCTGGTC	CTCGAGCGGC	540
CTTTGCCCGC	CCAGGATCTC	TTTGACTATA	TCACAGAGAA	GGGCCCACTG	GGTGAAGGCC	600
CAAGCCCGCT	CTTCTTTGGC	CAAGTAGTGG	CAGCCATCCA	GCACTGCCAT	TCCCGTGGAG	660
TTGTCCATGC	TGACATCAAG	GATGAGAACA	TCCTGATAGA	CCTACGCCGT	GGCTGTGCCA	720
AATCATTGA	TTTTGGTTCT	GGTGCCCTGC	TTCTGATAGA	ACCCTACACT	GACTTTGATG	780
GGACAAGGGT	GTACAGCCCC	CCAGAGTGGG	TCTCTCGACA	CCAGTACCAT	GCACTCCCGG	840
CCACTGTCTG	GTCACTGGGC	ATCCTCTCTT	ATGACATGGT	GTGTGGGGAC	ATTCCCTTTG	900
AGAGGGACCA	GGAGATTCTG	GAAGCTGAGC	TCCACTTCCC	AGCCCATGTC	TCCCCAGACT	960
GCTGTGCCCT	AATCCGCGCG	TGCTTGGCCC	CCAAACCTTC	TTCCGACCC	TCACTGGAAG	1020
AGATCTCTGT	GGACCCCTGG	ATGCAAAAC	CAGCCGAGGA	TGTTACCCCT	CAACCCCTCC	1080
AAAGGAGGGC	CTGCCCCCTT	GGCCTGTGTC	TTGTACCCCT	AAGCCTGGCC	TGGCCTGGCC	1140
TGGCCCCCAA	TGGTCAGAA	AGCCATCCCA	TGGCCATGTC	ACAGGGATAG	ATGACATTG	1200
GTTGACTTGG	TTTTACAGGT	CATTACCACT	CATTAAAGTC	CAGTATTACT	AAGGTAAGGG	1260
ATTGAGGATC	AGGGGTTAGA	AGACATAAAC	CAAGTTTGCC	CAGTTCCTCT	CCCAATCCTA	1320
CAAAGGAGCC	TTCTCCCGAG	AACCTGTGGT	CCCTGATTTT	GGAGGGGAA	CTTCTTGCTT	1380
CTCATTTTGC	TAAGGAAGTT	TATTTTGGTG	AAGTTGTTC	CATTTTGAGC	CCCGGGACTC	1440
TTATTTTGTG	GATGTGTGAC	CCCACATGG	CACCTCCTAC	TACCACCACA	CAAACTTAGT	1500
TCATATGCTT	TTACTTGGGC	AAGGGTGCTT	TCCTTCCAA	ACCCAGTAG	CTTTTATTTT	1560
AGTAAAGGGA	CCCTTTCCCG	TAGCCTAGGG	TCCCATATTG	GGTCAAGCTG	CTTACCTGCC	1620
TCAGCCGAGG	ATTTTATTTT	TTGGGGGAGG	TAATGCCCTG	TTGTTACCCC	AAGGCTTCTT	1680
TTTTTTTTTT	TTTTTTTTTT	GGTGAGGGGA	CCCTACTTTG	TTATCCCAAG	TGCTCTTATT	1740
CTGGTGAGAA	GAACCTTAAT	TCCATAATTT	GGGAAGGAAT	GGAAGATGGA	CACACCGGAA	1800
CACCACGAGA	CAATAGGATG	GGATGGATGG	TTTTTTGGGG	GATGGGCTAG	GGGAAATAAG	1860
GCTTGCTGTT	TGTTTCTCTG	GGGCGCTCCC	TCCAATTTTG	CAGATTTTGG	CAACCTCCTC	1920
CTGAGCCGGG	ATTGTCCCAAT	TACTAAATG	TAAATAATCA	CGTATTGTGG	GGAGGGGAGT	1980
TCCAAGTGTG	CCCTCTCTTT	TTTTCTGTCC	TGGATTATTT	AAAAAGCCAT	GTGTGGAAAC	2040
CCACTATTTA	ATAAAGTAA	TAGATCAGA	AAAAA	AAAAA		

Seq ID NO: 170 Protein sequence:
Protein Accession #: NP_006866

60
65
70

1	11	21	31	41	51	
MLTKPLQGP	P	PPGTPTPPP	GKDRFEA	EYRLGPLLGK	GGFGTVFAGH	RLTDLQVAI
KVIPRNRVLG	W	SPLSDSVTC	PLEVALLWKV	GAGGGHPGVI	RLLDWFETQE	GFMLVLERPL
PAQDLFDYIT	E	KGLPBGPS	RCFFGQVVA	IQHCHSRGVV	HRDIKDNIL	IDLRGCAKL
IDFGSGALLH	D	EPTDFDGT	RVYSPPEWIS	RHQYHALPAT	VWSLGLLYD	MVCGDIPFER
DQEILEAEHL	F	PARVSPDCC	ALIRRCAPK	PSSRPSLEEI	LIDPMNQTPA	EDVTPQPIQR
RPCFPGLVLA	T	LSLAWPLGA	PNGQKSFMA	MSQG		

Seq ID NO: 171 DNA sequence
Nucleic Acid Accession #: NM_003646
Coding sequence: 89..2875

75
80
85

1	11	21	31	41	51	
GCGGCGCGGA	GCGGCGGTGC	TGAGCCCCGG	CGCGCGGCC	GGCATGGGCG	TCTCCGCGG	60
GCCCTCGGCC	GGCGGGGGCT	AGGGCCGAT	GGAGCCGCG	GACGCTAGCC	CCGAGGCCCG	120
GAGCAGCGAC	TCCGAGTCGG	CTTCCGCTC	GTCCAGCGGC	TCCGAGCGCG	ACGCGGGTCC	180
CGAGCCGAGC	AAGCGCGCG	GGCGACTCAA	CAAGCGGCG	TTCCGCGGCG	TGCGGCTCTT	240
CGGGCAGAG	AAAGCCATCA	CCAAGTCGGG	CCTCCAGCAC	CTGGCCCCCG	CTCCGCCAC	300
CCCTGGGAGC	CGTGCAGCG	AGTCAGAGCG	GCAGATCCGG	AGTACAGTGG	ACTGGAGCGA	360
GTCAGCGACA	TATGGGGAGC	ACATCTGGTT	CGAGACCAAC	GTGTCGGGG	ACTTCTGCTA	420
CGTTGGGGAG	CAGTACTGTG	TAGCCAGGAT	GCTGAAGTCA	GTGCTCGAA	GAAAGTGCBC	480
AGCCTGCAAG	ATTGTGTGTC	ACACGCCCTG	CATCGAGCAG	CTGGAGAAGA	TAAATTTCCG	540
CTGTAAAGCG	TCCTTCCGTG	AATCAGGCTC	CAGGAATGTC	CGCGAGCCAA	CCTTTGTACG	600
GCACCACTGG	GTACACAGAC	GACGCCAGGA	CGGCAAGTGT	CGGCACTGTG	GGAAGGGATT	660
CCAGCAGAAG	TTCACTTCC	ACAGCAAGGA	GATTGTGGCC	ATCAGCTGCT	CGTGGTGCAA	720

5
10
15
20
25
30
35
40

```

GCAGGCATAC CACAGCAAGG TGTCTGCTT CATGCTGCAG CAGATCGAGG AGCCGTGCTC 780
GCTGGGGGTC CAGCGAGCGG TGGTCATCCC GCCCACCTGG ATCCTCCGCG CCCGGAGGCC 840
CCAGAAATCT CTGAAAGCAA GCAAGAAGAA GAAGAGGGCA TCCTTCAAGA GGAAGTCCAG 900
CAAGAAAGGG CTTGAGGAGG GCGCTGGAG ACCCTTCATC ATCAGGCCCA CCCCTCCCCC 960
GCTCATGAAG CCCCTGCTGG TGTTTGTGAA CCCCAAGAGT GGGGGCAACC AGGGTGCAAA 1020
GATCATCCAG TCTTTCTCTT GGTATCTCAA TCCCAGACAA GTCTTCGACC TGAGCCAGGG 1080
AGGGCCCAAG GAGCGCTGG AGATGTACCG CAAAGTGCAC AACCTGCGGA TCCTGGCGTG 1140
CGGGGGCGAC GGCAGCGTGG GCTGGATCCT CTCCAACCTG GACCAGTAC GCCTGAAGCC 1200
GCCACCCCTT GTTGCCATCC TGCCCCCTGG TACTGGCAAC GACTTGGCCC GAACCTCAA 1260
CTGGGGTGGG GGCTACACAG ATGAGCCTGT GTCCAAGATC CTCTCCACG TGGAGGAGGG 1320
GAACGTGGTA CAGCTGGACC GCTGGGACCT CCACGCTGAG CCCAACCCG AGGCAGGGCC 1380
TGAGGACCGA GATGAAGGCG CCACCGACCG GTTGCCCTTG GATGTCTCA ACAACTACTT 1440
CAGCCTGGGG TTTGACGCCC AOGTCAACCT GGAGTTCAC GAGTCTCGAG AGGCCAACCC 1500
AGAGAAATTC AACAGCCGCT TCGGAATAA GATGTTCTAC GCGGGACAG CTTTCTCTGA 1560
CTTCTGATG GGCAGCTCCA AGGACCTGGC CAAGCACATC CGAGTGGTGT GTGATGGAAT 1620
GGACTTGACT CCCAAGATCC AGGACCTGAA ACCCCAGTGT GTTGTCTTCC TGAACATCCC 1680
CAGGTAAGT GCGGGCAGCA TGCCCTGGGG CCACCTGGG GAGCACCAGC ACTTTGAGCC 1740
CCAGGGCAT GAGCAGCGCT ACCTCGAGGT CATTTGGCTC ACCATGACGT CGTTGGCCCG 1800
GCTGCAGGTG GCGGACAGC GCGAGCGGCT GACGAGTGT CCGAGGTGG TGCTCACAC 1860
ATCCAAGGCC ATCCCGGTGC AGGTGGATGG CGAGCCCTGC AAGCTTGAG CCTCACGCAT 1920
CGGCATCGCC CTGCGCAACC AGGCCACCAT GGTGCAGAG GCCAAGCGC GGCAGCGCCG 1980
CCCCCTGAC AGCGACAGC AGCCGGTGCC AGAGCAGTTG GGCATCCAG TGAGTCCGCT 2040
CAGCATGCAC GACTATGAGG CCCTGCACTA CGACAAGGAG CAGCTCAAG AGGCCTCTGT 2100
GCGCTGGGG ACTGTGGTGG TCCCAGGAGA CAGTGACCTA GAGCTCTGCC GTGCCACAT 2160
TGAGAGACTC CAGCAGGAGC CCGATGGTGC TGGAGCCAA TCCCGACAT GCCAGAACT 2220
GTCCCCAGAG TGGTGTCTCC TGGACGCCAC CACTGCCAGC CGCTTCTACA GGATCGACCG 2280
AGCCAGAGG CACTCAACT ATGTGACTGA GATCGCACAG GATGAGATT ATATCTGGA 2340
CCCTGAGCTG CTGGGGGCAT CGGCCCGGCC TGACCTCCCA ACCCCACTT CCCCTCTCCC 2400
CACCTCACCC TGCTCACCCA CGCCCGGTC ACTGCAAGGG GATGCTGCAC CCCCTCAAG 2460
TGAAGAGCTC ATTTGAGGCT CCAAGAGGAA CGACTTCTGT AAGCTCCAGG AGCTGCACCG 2520
AGCTGGGGGC GACTCATGAC ACCGAGACGA GCAGAGTCCG ACGCTCTGCG ACCACGAGT 2580
CAGCACTGCG AGCAAGGATG TGGTCCGCTA CCTGCTGGAC CACGCCCCCG CAGAGATCCT 2640
TGATGCGGTG GAGGAAAACG GGGAGACCTG TTTGCACCAA GCAGCGGCC TGGGCCAGCG 2700
CACCATCTCG CACTACATCG TGGAGGCCGG GGCCTCGCTC ATGAAGACAG ACCAGCAGGG 2760
CGACACTCCC CGCAGCGCGG CTGAGAAGGC TCAGGACACC GAGCTGGCCG CCTACCTGGA 2820
GAACCGGCAG CACTACCAGA TGATCCAGCG GGAGGACCA GAGACGGCTG TGTAGCGGGC 2880

```

Seq ID NO: 172 Protein sequence:
Protein Accession #: NP_003637

45
50
55
60

```

1 11 21 31 41 51
MEPRDGSPEA RSSDSSESASA SSSGSESDAG PEPDKAPRRL NKRRPPGLRL FGRKAITKS 60
GLQHLAPPPP TPGAPCSESE RQIRSTVDWS ESATYGEHIW PETNVSGDPC YVGEQYCVAR 120
MLKSVSRKRC AACKIVVHTP CIEQLEKINF RCKPSFRESG SRNVREPTFV RHHVVERRRQ 180
DGKCRHCGKG FQCKTFPHSK EIVAISCSCW QKAYHSKVSC FMLQIEEPC SLGVHAAVVI 240
PPTWILRAR R PONTLKASKK KKRASFRRKS SKKGPBERGRW RFFIIRPTPS PLMKPLLVFV 300
NPKSGGNQGA KIIQSPLWYL NPRQVFDLSQ GGPKEALEMY RKVHNLRLA CGGDTGVGI 360
LSTLDLRLK PPPFVAILPL GTGNDLARTL NWGGYTDPE VSKILSHVEE GNVVQLDRWD 420
LHAEPNPEAG PEDRDEGATD RLPLDVFNRY FSLGFDHVT LEPHESREAN PEKFNFRFN 480
KMFYAGTAFS DFLMGSKDL AKHIRVVDG MDLTPKIQDL KPQCVVPLNI PRYAGTMPW 540
GHPGEHDFE PQRHDGYLE VIGFTMTSLA ALQVGGHGER LTQCREVLT TSKAIPQVD 600
GEPCKLAASR IRIALRNQAT MVQKAKRSA APLHSDQQPV PEQLRIQVSR VSMHDEYALH 660
YDKBQLKEAS VPLGTVVVPG DSDLELCRAH IERLQEPDG AGAKSPCTCK LSPKWCFLDA 720
TTASRFYRID RAGEHLYVT EIAQDEIYIL DPELLGASAR PDLPTPSPL PTSPCSPTR 780
SLOGDAAPQ GEELIEAAKR NDFCKLQELH RAGGDLMRD EQSRTLHHA VSTGSKDVVR 840
YLLDHAPPEI LDAVENGET CLHQAAALGQ RTICHYIVEA GASLMKTDQ GDTPRQRAEK 900
AQDTLAAYL ENRQHYMIQ REDQETAV

```

Seq ID NO: 173 DNA sequence
Nucleic Acid Accession #: AF232772
Coding sequence: 1-1662

65
70
75
80
85

```

1 11 21 31 41 51
ATGCGGGTGC AGCTGACGAC AGCCCTGCGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60
GTGCTGGGTG GCATCTCTGC AGCCTATGTG ACGGGCTACC AGTTCAATCCA CACGGAAG 120
CACTACCTGT CCTTCGGCCT GTACGGCGCC ATCCTGGGCC TGCACTTGCT CATTCAGAGC 180
CTTTTTCGCT TCCTGGAGCA CGGCGCATG CGACGTGCGC GCCAGGCCCT GAACTGCC 240
TCCCGCGGCG GGGCTCGTGT GGCATGTGCT ATTGCGCAT ACCAGGAGGA CCCTGACTAC 300
TTGCGCAAGT GCTTCGCTC GGCACAGCGC ATCTCCTTCC CTGACCTCAA GGTGGTCATG 360
GTGGTGGATG GCAACCGCCA GAGGACGCGC TACATGCTGG ACATCTTCCA CGAGGTGCTG 420
GGCGGCACCG AGCAGGCGCG CTTCTTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG 480
GGTGAGACCG AGGCCAGCCT GCAGGAGGGC ATGACCGGTG TCGGGGATGT GGTGCGGGCC 540
AGCACCTTCT CGTGCAATCAT GCAGAAGTGG GGAGGCAAGC GCGAGGTCTA GTACACGGCC 600
TTCAAGGCCCT TCGCGGATTC GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660
GATCCAGCCT GCACCATCGA GATGCTTCCA GTCTTGGAG AGGATCCCCA AGTAGGGGGA 720
GTGGGGGAG ATGTCCAGAT CCTCAACAG TACGACTCAT GGATTTCCTT CCTGAGCAGC 780
GTGGGTACT GGAATGCCCTT CAACGTGGAG CGGGCCTGCC AGTCTACTT TGGCTGTGTG 840
CAGTGTATTA GTGGGCCCTT GGCATGTAT CCGCAACAGC TCCTCCAGCA GTTCTGGAG 900
GACTGTATAC ATCAGAAGTT CCTAGGCAGC AAGTGACGCT TCGGGGATGA CCGCACCTC 960
ACCAACCGAG TCCTGAGCCTT TGGCTACCGA ACTAAGTATA CCGCGCGCTC CAAGTGCCCT 1020
ACAGAGACCG CCACTAAGTA CCTCGGTGG CTCAACGAGC AAACCCGCTG GAGCAAGTCT 1080
TACTTCCGGG AGTGGCTCTA CAACCTCTCT TGGTTCCATA AGCACCACCT CTGATGACC 1140
TACGAGTCAG TGGTACCGGG TTTCTTCCCC TTCTTCTCTA TTGCCACGGT TATACAGCTT 1200
TTCTACCGGG GCGGCATCTG GAACATTCTC CTCTTCTCTG TGACGGTGCA GCTGGTGGGC 1260
ATTATCAAGG CCACCTACGC CTGCTTCTCT CGGGGCAATG CAGAGATGAT CTTCTATGCC 1320

```

	CTCTACTCCC	TCCTCTATAT	GTCCAGCCTT	CTGCCGCGCA	AGATCTTTCG	CATTGCTACC	1380
	ATCAACAAAT	CTGGCTGGGG	CACCTCTGGC	CGAAAAACCA	TTGTGGTGAA	CTTCATTGGC	1440
	CTCATTCTCG	TGTCATCTCG	GGTGGCAGTT	CTCCTGGAGG	GGCTGGCCTA	CACAGCTTAT	1500
	TGCCAGGACC	TGTTAGTGA	GACAGAGCTA	GCCTTCCTTG	TCTCTGGGGC	TATACTGTAT	1560
5	GGCTGCTACT	GGGTGGCCCT	CCTCATGCTA	TATCTGGCCA	TCATCGCCCG	GCGATGTGGG	1620
	AAGAAGCCGG	AGCAGTACAG	CTTGGCTTTT	GCTGAGGTGT	GACATGGCCC	CCAAGCAGAG	1680
	CGGGTAAAGT	GCAATGGGTA	AGGGAGGGAA	GGGGAATGGA	AGAGAAAAGA	CAGGGTGGGA	1740
	GGGAGGAGGG	AGTGTCTGT	TTTAGTCTCT	TAATGGTCCA	AAGGACAAAT	CTAAATAGCA	1800
	AAGAAGCGTG	ATGTAGTATG	GCCTGACAGC	TCTGTTTAGA	GGAGGCAACA	CTGATCCCCC	1860
10	AGATGCAGGG	CTGCAGGGGA	TTCTGTGTTT	TCAGACTGCC	TGCTGTCTTG	CATCTGCACA	1920
	TAGGCAGTAG	CCTCCTCCTG	GGCTCCAGAG	GGCACTCAGA	AGTGTGTCTA	AACCAAGTTA	1980
	AGTCCCATTC	AGTGGCAACT	TGTGATAGGT	ACCTGAGTGA	CGGCAACCTG	CGGAAGGAGG	2040
	TTCTCCGACG	CCATCTGAAC	ACAACCAGAG	GTGGCAGGAG	AATTTCTACT	GAGCGAGGTG	2100
	GGCGGTTTAG	TGTATGTCAC	CCCCACCCCA	CCCATAGTA	GTCAATCATG	CAATAAGATT	2160
15	GCGCGTGAGA	TACAAGGCCC	AGAAGCCTGA	TCTTTGGGCA	TCAGAAAACA	GGGTCCAGGA	2220
	ATGGTGCCTT	ATGTGAGATA	CCCCACTCCA	CATCAACATT	CCAGGGATGA	GCCAAACAGG	2280
	CAGGGAGTGA	GCACTGGAAT	GCTTTTAAAA	GTGCACATTA	AAAAGGAAAG	TTTGCCAGGA	2340
	GGAAACAAAG	GATTGTGGTG	GTGCTAAAGG	AGGCCATAAG	CTACACAGAG	GCCTTGGGTG	2400
	TTCCACCTGG	AAACTGCTCA	GACGTCTAGA	TGGGTTCTTA	GCTGTCTGT	GATCTCTGCT	2460
20	GGGAGATAAA	AAAGATTAA	CCCCAACATG	TTGAGAAAAG	AAGTGAAGTC	TTGGGTATTT	2520
	TAACTGTAT	ACTCTTGAAT	TCCTCTCAAA	TTGAGCTCTG	ATCTGAGGCT	AAGACACACT	2580
	CCCACTTCA	CTTCTTCA	AGCCACATTT	TTTGAGGTAT	CACTGCAGTC	ACCTCTTCTA	2640
	CCCTCATCAT	CATAGGTAAG	GTTTTCAGG	TGGCAATTGG	GGCGGAGCCC	CGGCTTCTTA	2700
	TAGAAGCTTC	AGCAGGAGGC	AAGCGTGTTC	TCAGCACATA	TGGGAAGTAT	GAGGAGCCTC	2760
25	TGATCAAAAT	GGCTACAATC	TTGGAGCTGC	TTGGACGGAT	TCCTTGGCAG	CGGGGTTAGC	2820
	ATGTGTGACT	TTCAAGGCTAC	TGTTCTTGAC	AATCATCTCC	AATGGAAGC	TTTTTCAGTG	2880
	TCCCAAGTGA	AACTCTCAAA	TCCAAAATGG	TTATCTTTGA	GACCATCCAT	TCTCCTCAGT	2940
	GGCTTCTCCA	GGGAATCTCT	ACAGCCAAAGT	TGTGACAGTC	ACTGCATTTG	CCTGCTTCTT	3000
	TCCAGAAACC	AAACTAGGAG	ATGAACTGG	TTCTACATC	CTAAGGTTCT	TGCTTCTCT	3060
30	CTCATGCCCT	CTGAGGCTGT	TTTGGCTGT	TTTCCCTCTG	CTGCTTTTGG	GGAATGAGGG	3120
	GAAGCCATT	TCCAAGTGAC	TTGCAATCCA	GGCTGTCTCT	AGCGTTTGA	GTTTAAACCC	3180
	TGGGATCCCTG	CTAAGGCTTT	TGACTTAAGG	GTGCTGTCT	TGCCCTCCAA	ATGCTCTTTC	3240
	TCAAAGGGGC	CAACTAACCC	GTGCAGAAC	AGCACTAAGG	TGGACAGCAG	ACAAGAGGGC	3300
	AAGCCTCTAA	TGTACCAAGT	GCTTCTCTACA	AAGACGCAAG	GTGTGCTCCG	AACCACAGAT	3360
35	GGGCAAAACC	TGGTGTCTTC	CTTCACTCTC	CACGAACTCA	AGGGTTTCCG	AAGTGTAGCT	3420
	AACAGTTGCC	ACATCACACA	GACCTCCAGT	TTCTGGTAAG	ACTGCTGGTT	GACATCAGAC	3480
	CCAAACCAAT	GAAGGCTGGA	AGGCAGCAGG	CATTGCTTAA	GGCAGCTGAT	CCAGGCAATC	3540
	GTTCTGCTGG	CCAAGAAGTT	AAACTATTTT	GAGCATTAGA	ATGGAGGAAA	TCCGGTCAGC	3600
	CAAGTGCAGA	GTTGAGACT	CGCTAAGGGC	TTGTTTTTCT	TCAGCATTTA	CTTGAAGATT	3660
40	AATGTAGGAT	GACAGGCTCT	CCTGGCTGTC	CTACCATCAG	CTCTGCCTTG	CACCTGTGTC	3720
	GTCACACTTC	CTCAAACTCA	AAACAGGCAG	GTACAGGTAG	TGGGCTCACA	ACGTTTGACC	3780
	TGCACTGGTT	TTTCTAAGTT	ATTTTGTACA	TTTTTCAGCA	GCAAAACCAA	ACTGGGTCTT	3840
	CAGCTTTATC	CCCGTTTCTT	GCAAGGGAAG	AGCCTTTATA	CAATTGGACG	CATTTTGGTT	3900
45	TTTCTCTCAT	GAGAATTCAA	ATCCTCTTTT	GTATTGTTC	TACAATAATT	TGTAACATA	3960
	TTTATTTTA	CCTGCTTTT	TTTTTTTTT	TAATTTTCAG	GTCAAGTTTT	TTATACTGCA	4020
	CTTATTGTG	AAAATAAGA	TTCTCACAT				

Seq ID NO: 174 Protein sequence:
Protein Accession #: AAF36984

50	1	11	21	31	41	51	
	MPVQLTTALR	VVGTSFLPALA	VLGGLIAAYV	TGYQFIHTEK	HYLSFGLYGA	ILGLHLLIQS	60
	LFAFLERRRM	RRAGQALKLP	SPRRGSVALC	IAAYQEDPDY	LRKCLRSAGR	ISFPDLKVVM	120
55	VVDGNRQEDA	YMLDIPHEVL	GGTEQAGFFV	WRSNFHEAGE	GETEASLQEG	MDRVDRVVRA	180
	STFSCIMQKW	GGKREVMYTA	FKALGDSVDY	IQVCDSDIVL	DPACTIEMLR	VLEEDPQVGG	240
	VGGDVQILNK	YDSWISFLSS	VRYWMAFNVE	RACQSYFGCV	QCISGFLGMY	RNSLLQQLFL	300
	DWYHQKFLGS	KCSFGDDRHL	TNRVLSLGYR	TKYTARSKCL	TETPTKYLRW	LNQQRWSKS	360
	YPREWLXNSL	WFHKLHLMWT	YESVVTGFFP	FPLIATVIQL	FYRGRIWNIL	LFLLTVQLVGL	420
60	LIKATYACFL	RGNAMIFMS	LYSLLYMSSL	LPAKIPAIAT	INKSGWGTSG	RKTIIVNFIG	480
	LIPVSIWVAV	LLEGLAYTAY	CQDLFSETEL	APLVSGAILY	GCYWVALMLL	YLAIIARRCG	540
	KKPQYSLAF	AEV					

Seq ID NO: 175 DNA sequence
Nucleic Acid Accession #: NM_000691
Coding sequence: 43..1404

70	1	11	21	31	41	51	
	CCAGGAGCCC	CAGTTACCGG	GAGAGGCTGT	GTCAAAGGGG	CCATGAGCAA	GATCAGCGAG	60
	GCGGTGAAGC	GCGCCCGCGC	CGCCTTCAGC	TGCGGCAGGA	CCCGTCCGCT	GCACTTCCGA	120
	TTCCAGCAGC	TGGAGGCGCT	GCAGGCGCTG	ATCCAGGAGC	AGGAGCAGGA	GCTGGTGGGC	180
	GCGCTGGCCG	CAGACCTGCA	CAAGAATGAA	TGGAAGCGCT	ACTATGAGGA	GGTGGTGTAC	240
75	GTCCTAGAGG	AGATCGAGTA	CATGATCCAG	AAGCTCCCTG	AGTGGGCGGC	GGATGAGCCC	300
	GTGGAGAGA	CGCCCCAGAC	TCAGCAGGAC	GAGCTCTACA	TCCACTCGGA	GCCACTGGGC	360
	GTGGTCTCTG	TCATTGGCAC	CTGGAACTAC	CCCTTCAACC	TCACCATCCA	CCCCATGGTG	420
	GGCGCCATGC	GATCAGGGAA	CGCAGTGGTC	CTCAAGCCCT	CGGAGCTGAG	TGAGAACATG	480
	GCGAGCGCTG	TGGCTACCAT	CATCCCCCAG	TACCTGGACA	AGGATCTGTA	CCCAAGTAATC	540
80	AATGGGGCTG	TCCCTAGAGC	CACGGAGCTG	CTCAAGGAGA	GGTTGACCCA	TATCCTGTAC	600
	ACGGGCAGCA	CGGGGGTGGG	GAAGATCATC	ATGACGGCTG	CTGCCAAGCA	CCTGACCCCT	660
	GTCACGCTGG	AGCTGGGAGG	GAAGAGTCCC	TGCTACGTGG	ACAAGAAGTG	TGACCTGGAC	720
	GTGGCCTGCC	GATCGATCGC	CTGGGGGAAA	TTCATGAACA	GTGGCCAGAC	CTGCGTGGCC	780
	CCAGACTACA	TCTCTGTGTA	CCCTCGATC	CAGAAACCAA	TTGTGGAGAA	GCTCAAGAAG	840
85	TACTGTAAAG	AGTTCTACGG	GGAAGATGCT	AAGAAATCCC	GGGACTATGG	AAGAATCATT	900
	AGTGCCCGGC	ACTTCCAGAG	GGTGATGGGC	CTGATTGAGG	GCCAGAAGGT	GGCTTATGGG	960
	GGCACCGGGG	ATGCGCGCAC	TGCTACATA	GCCCCACCA	TCTTCACGGA	CGTGGACCCC	1020

5
10
15
20
25

```

CAGTCCCCG  TGATGCAAGA  GGAGATCTTC  GGGCCTGTGC  TGCCCATCGT  GTGCGTGGCC  1080
AGCCTGGAGG  AGGCCATCCA  GTTCATCAAC  CAGCGTGAGA  AGCCCCGGCC  CCTCTACATG  1140
TTCTCCAGCA  AGCAACAGGT  GATTAAGAAG  ATGATTGCAG  AGACATCCAG  TGGTGGGGTG  1200
GCGGCCAAGC  ATGTATCGT  CCACATCACC  TTGCACTCTC  TGCCCTTCGG  GGGCGTGGGG  1260
AACAGCGGCA  TGGGATCCTA  CCATGGCAAG  AAGAGCTTCG  AGACTTTCCT  TCACCGCGCC  1320
TCTTGCTCG  TGAGGCTCT  GATGAATGAT  GAAGGCTCGA  AGGTGAGATA  CCCCCGAGC  1380
CCGCGCAAGA  TGACCCAGCA  CTGAGGAGGG  GTTGCTCCGC  CTGGCTGGCC  CATACTGTGT  1440
CCCATCGGAG  TCGGAGCCAC  CCTCACTGGC  TCTCTGGCCC  CTGGAGAATC  GCTCTGCAG  1500
CCCCAGCCCA  GCCCCATCC  TCTGCTGACC  TGCTGACCTG  TGCACACCCC  ACTCCACAT  1560
GGGCCAGGC  CTCACATTC  CAAGTCTCCA  CCCCCTTTCTA  GACCAATAAA  GAGACAAATA  1620
CAATTTTCTA  ACTCGG

```

Seq ID NO: 176 Protein sequence:
Protein Accession #: NP_000682

15
20
25

```

1      11      21      31      41      51
MSKISEAVKR  ARAAFSSGR  RPLQFRPQOL  EALQRLIQEQ  EQELVGALAA  DLHKIENWY  60
YEEVVVYLEE  IEYMIQKLE  WAADEPVEKT  PQTQDELYI  HSEPLGVVLV  IGTWYNNPL  120
TIQPMVGAIA  AGNAVVLKPS  ELSENMASLL  ATIIQYLDK  DLYFVINGGV  PETTELLKER  180
FDHILYTGST  GVGKILMTAA  AKHLTPVTLE  LGGKSPCYVD  KNCDLDVACR  RIANGKFMNS  240
GQTCVADPYI  LCDPSIQNQI  VEKLIKSLKE  FYGEDAKKSR  DYGRISARH  FQRVMGLIEG  300
QKVAYGGTGD  AATRYLAPT  LTDVDPQSPV  MQEIEFGFVL  PIVCVRSLEE  AIQPINQREK  360
PLALYMFSSN  DKVIKKMAE  TSSGVAAND  VIVHITLHSL  PFGGVNSGM  GSYHGKKSFE  420
TFSHRSLV  RFLMDEGLK  VRYPPSPAKM  TQH

```

Seq ID NO: 177 DNA sequence
Nucleic Acid Accession #: NM_001067.1
Coding sequence: 108-4703

30
35
40
45
50
55
60
65
70
75
80
85

```

1      11      21      31      41      51
CTAACCGACG  CGGCTCTGTG  GAGAAGCGGC  TTGGTCGGGG  GTGGTCTCGT  GGGGTCTGTC  60
CTGTTTAGTC  GCTTTCAGGG  TTCTTGAGCC  CCTTCACGAC  CGTCACCATG  GAAGTGTCTAC  120
CATTCGAGCC  TGTAAATGAA  AATATGCAAG  TCAACAAAT  AAAGAAAAAT  GAAGATGCTA  180
AGAAAAAGACT  GCTGTGTGAA  AGAATCTATC  AAAAGAAAA  ACAATGGGAA  CATATTTTGC  240
TCGCCCCAGA  CACCTACATT  GGTCTGTGG  AATTAGTGAC  CCAGCAAAATG  TGGGTTTACG  300
ATGAAGATGT  TGGCATTAA  TATAGGGAAG  TCACITTTGT  TCCTGGTTTG  TACAAATCT  360
TTGATGAGAT  TCTAGTTAAT  GCTGCGGACA  ACAACAAAG  GGACCCAAAA  ATGCTCTGTA  420
TTAGAGTCAC  AATTGATCCG  GAAAACAATT  TAATTAGTAT  ATGGAATAAT  GAAAAAGGTA  480
TTCTGTGTGT  TGAACACAAA  GTTGAAGAAG  TGTATGTCCC  AGCTCTCATA  TTTGGACAGC  540
TCCTAACTTC  TAGTAACAT  GATGATGATG  AAAAGAAAGT  GACAGGTGGT  CGAAATGGCT  600
ATGGAGCCAA  ATTGTGTAAC  ATATTCAGTA  CCAAAATTAC  TGTGGAACA  GCCAGTAGAG  660
AATACAAGAA  AATGTTCAA  CAGACATGGA  TGGATAATAT  GGAAGAGCT  GGTGAGATGG  720
AACTCAAGCC  CTTCAATGGA  GAAGATTATA  CATGTATCAC  CTTTCAGCCT  GATTGTCTA  780
AGTTTAAAT  GCAAAGCCGT  GACAAAGATA  TTGTGCACT  AATGGTCAGA  AGAGCATATG  840
ATATTGCTGG  ATCCACCAA  GATGTCAAAG  TCTTCTTAA  TGGAAATAAA  CTGCCAGTAA  900
AAGGATTTCG  TAGTTATGTG  GACATGTATT  TGAAGGACAA  GTTGGATGAA  ACTGGTAACT  960
CCTTGAAAGT  AATAGTCAA  CAAGTAAACC  ACAGGTGGGA  AGTGTGTTTA  ACTATGAGTG  1020
AAAAAGCTT  TCAGCAAAAT  AGCTTTGTCA  ACAGCAATGC  TACATCCAAG  GGTGGCAGAG  1080
ATGTTGATTA  TGTAGCTGAT  CAGATTGTGA  CTAACTTGT  TGATGTTGTG  AAGAAGAAGA  1140
ACAAGGGTGG  TGTTCAGTA  AAAGCACATC  AGGTGAAAA  TCACATGTGG  ATTTTGTAA  1200
ATGCCCTAAT  TGAAACCCA  ACCTTTGACT  CTCAGACAAA  AGAAAAACATG  ACTTTACAAC  1260
CAGAGAGCTT  TGGATCAACA  TGCCAATTGA  GTGAAAAATT  TATCAAAGCT  GCCATTGGCT  1320
GTGGTATTGT  TCAAGACATA  CTAACTGGG  TGAAGTTTAA  GGCCCAAGTC  CAGTTAAACA  1380
AGAAGTGTTC  AGCTGTAAA  CTAATAGAA  TCAAGGGAAT  TCCCAAATC  GATGATGCCA  1440
ATGATGCAGG  GGGCCGAAAC  TCCACTGAGT  GTACGCTTAT  CCTGACTGAG  GGAGATTGAG  1500
CCAAACTTT  GCGTGTTC  GGCCTTGGTG  TGGTTGGGAG  AGACAAATAT  GGGGTTTTCC  1560
CTCTTAGAG  AAAAATCTC  AATGTTGAG  AAGCTTCTCA  TAAGCAGATC  ATGGAAAAATG  1620
CTGAGATTAA  CATATCATC  AAGATTGTGG  GTCTTCAGTA  CAAGAAAAAC  TATGAAGATG  1680
AAGATTCAAT  GAAGACGCTT  CGTTATGGGA  AGATAATGAT  TATGACAGAT  CAGGACCAAG  1740
ATGGTTCCCA  CATCAAAGGC  TTGCTGATTA  ATTTTATCCA  TCACAATG  CCGTCTCTTC  1800
TGCGACATCG  TTTTCTGGAG  GAATTTATCA  CTCCATTGT  AAAGGTATCT  AAAACAAGC  1860
AAGAAATGGC  ATTTTACAGC  CTTCCTGAAT  TTGAAGAGTG  GAAGAGTTCT  ACTCCAATC  1920
ATAAAAAATG  GAAAGTCAAA  TATTACAAG  GTTGGGCAC  CAGCACATCA  AAGGAAGCTA  1980
AAGAATACTT  TGCAGATATG  AAAAGACATC  GTATCCAGTT  CAAATATTCT  GGTCTGAAG  2040
ATGATGCTGC  TATCAGCCTG  GCCTTTAGCA  AAAACAGAT  AGATGATCGA  AAGGAATGGT  2100
TAACTAATTT  CATGAGGAT  AGAAGACAA  GAAAGTTACT  TGGGCTTCT  GAGGATTACT  2160
TGTATGGAAC  AACTACCA  TATCTGACAT  ATAATGACTT  CATCAACAAG  GAACTTATCT  2220
TGTCTCAAA  TTCTGATAAC  GAGAGATCTA  TCCCTTCTAT  GGTGGATGGT  TTGAAACCAG  2280
GTCAGAGAAA  GGTTTTGT  ACTTGCTTCA  AACGGAATGA  CAAGCGAGAA  GTAAAGGTG  2340
CCCAATTAGC  TGGATCAGTG  GCTGAAATGT  CTCTTATCA  TCATGGTGAG  ATGTCATAA  2400
TGATGACCAT  TATCAATTG  GCTCAGAATT  TTGTGGGTAG  CAATAATCTA  AACCTCTG  2460
AGCCCATTTG  TCAGTTTGGT  ACCAGGCTAC  ATGGTGGCAA  GGATCTGCT  AGTCCAGAT  2520
ACATCTTTAC  AATGCTCAGC  TCTTGGCTC  GATTGTTATT  TCCACCAAAA  GATGATCACA  2580
CGTTGAAGTT  TTTATATGAT  GACAACGAC  GTGTTGAGCC  TGAATGGTAC  ATTCCTATTA  2640
TTCCCATGGT  GCTGATAAAT  GGTGCTGAAG  GAATCGGTAC  TGGGTGGTCC  TGCAAAATCC  2700
CCAACCTTGA  TGTGTGGAAG  ATGACTGAAG  AAAAAGTGGC  AGAGGCAGAG  AGAGTTGGAC  2760
CTTGCCAAAT  GCTTCAAGT  TACAAGAACT  TCAAGGGTAC  TATTGAAGAA  CTGGCTCCAA  2820
ATCAATATGT  GATTAGTGGT  GAAGTAGCTA  TTCTTAATTC  TACAACCAAT  GAAATCTCAG  2880
AGCTTCCCGT  AAGACATGG  ACCCAGACAT  ACAAGAACA  AGTTCTAGAA  CCCATGTTGA  2940
ATGGCACCGA  GAAGACACCT  CCTCTATAA  CAGACTATAG  GGAATACCAT  ACAGATACCA  3000
CTGTGAATTT  TGTGTGGAAG  ATGACTGAAG  AAAAAGTGGC  AGAGGCAGAG  AGAGTTGGAC  3060
TACACAAAGT  CTCAAACTC  CAAACTAGTC  TCACATGCAA  CTCTATGGTG  CTTTTTGACC  3120
ACGTAGGCTG  TTTAAGAAA  TATGACACGG  TGTGGATAT  TCTAAGAGAC  TTTTTTGAAC  3180
TCAGACTTAA  ATATTATGGA  TTAAGAAAAG  AATGGCTCCT  AGGAATGCTT  GGTGCTGAAT  3240
CTGCTAAACT  GAATAATCAG  GCTCGCTTTA  TCTTAGAGAA  AATAGATGGC  AAAATAATCA  3300

```

	TTGAAATAA	GCCTAAGAAA	GAATTAATTA	AAGTCTGTAT	TCAGAGGGGA	TATGATTGGG	3360
	ATCCTGTGAA	GGCCTGGAAA	GAAGCCGAGC	AAAAGGTTCC	AGATGAAGAA	GAAAATGAAG	3420
	AGAGTGACAA	CGAAAAGGAA	ACTGAAAAGA	GTGACTCCGT	AACAGATTCT	GGACCAACCT	3480
5	TCAACTATCT	TCTTGATATG	CCCTTTGGT	ATTTAAACCA	GGAAAAGAAA	GATGAACCT	3540
	GCAGGCTAAG	AAATGAAAAA	GAACAAGAGC	TGGACACATT	AAAAAGAAAG	AGTCCATCAG	3600
	ATTTGTGGAA	AGAAGACTTG	GCTACATTTA	TTGAAGAATT	GGAGGCTGTT	GAAGCCAAGG	3660
	AAAAACAAG	TGAACAAGTG	GGACTTCCCT	GGAAAGGGGG	GAAGGCCAAG	GGGAAAAAAA	3720
	CACAAATGCG	TGAAGTTTGG	CCTTCTCCGC	GTGGTCAAAAG	AGTCATTCCA	CGAAATAACCA	3780
10	TAGAAATGAA	AGCAGAGGCA	GAAAAGAAAA	ATAAAAAGAA	AATTAAGAAT	GAAAATACTG	3840
	AAGGAAGCCC	TCAAGAAGAT	GGTGTGGAAC	TAGAAGGCCT	AAAAACAAGA	TTAGAAAAGA	3900
	AACAGAAAAG	AGAACCAGGT	ACAAAGACAA	AGAAACAAAC	TACATTGGCA	TTTAAGCCAA	3960
	TCAAAAAAGG	AAAGAAGAGA	AATCCCTGGC	CTGATTTCAGA	ATCAGATAGG	AGCAGTGACG	4020
	AAAGTAATTT	TGATGTCCCT	CCACGAGAAA	CAGAGCCACG	GAGAGCAGCA	ACAAAAACAA	4080
	AATTCACAAT	GGATTTCGGAT	TCAGATGAAG	ATTTCTCAGA	TTTTGATGAA	AAAACTGATG	4140
15	ATGAAGATTT	TGTCCCATCA	GATGCTAGTC	CACCTAAGAC	CAAACTTCC	CCAAAACTTA	4200
	GTAACAAAAG	ACTGAAACCA	CAGAAAAGTG	TGCTGTACAG	CCTTGAAGCT	GATGATGTTA	4260
	AGGGCAGTGT	ACCACTGTCT	TCAAGCCCTC	CTGCTACACA	TTTCCAGAT	GAAACTGAAA	4320
	TTACAAACCC	AGTTCCTAAA	AAGATGTGA	CAGTGAAGAA	GACAGCAGCA	AAAAGTCAGT	4380
20	CTTCCACCTC	CACCTACCGT	GCCAAAAAAA	GGGCTGCCCC	AAAAGGAAC	AAAAGGGATC	4440
	CAGCTTTGAA	TCTGGGTGTC	TCTCAAAAGC	CTGATCTGCG	CAAAACCAAG	AATCGCCGCA	4500
	AAAGGAAGCC	ATCCACTTCT	GATGATTCTG	ACTCTAATTT	TGAGAAAAAT	GTTCGAAAG	4560
	CAGTCAACAG	CAAGAATCC	AAGGGGGAGA	GTGATGACTT	CCATATGGAG	TTTGACTCAG	4620
	CTGTGGCTCC	TGGGGCAAAA	TCTGTACGGG	CAAGAAACCC	TATAAGTAC	CTGGAAGAGT	4680
	CAGATGAAGA	TGATCTGTTT	TAAAATGTGA	GGCGATTATT	TTAAGTAATT	ATCTTACCNA	4740
25	GCCCAAGACT	GGTTTAAAG	TTACCTGAAG	CTCTTAACCT	CCTCCCCCTC	GAATTTAGTT	4800
	TGGGGAGAGT	GTTCCTTAGT	CAAGACATCA	AAGTGAAGTA	AAGCCCAAGT	GTTCCTTAGC	4860
	TTTTTATAGT	AGATCTTAA	TAGTGACCAT	CTCATGGGCA	TTGTTTCTT	CTCTGCTTGG	4920
	TCTGTGTTT	GAGTCTGCTT	TCTTTTGTCT	TAAAAACCTG	ATTTTAAAGT	TCTCTGAAC	4980
	TGTAGAATA	GCTATCTGAT	CACCTCAGCG	TAAAGCAGTG	TGTTTATTAA	CCATCCACTA	5040
30	AGCTAAAAGT	AGAGCAGTTT	GATTTAAAG	TGTCACCTCT	CCTCCTTTTC	TACTTTAGTT	5100
	AGATATGAGA	TAGAGCATAA	TTATCTGTTT	TATCTTAGTT	TTATACATAA	TTTACCATCA	5160
	GATAGAACTT	TATGGTTCTA	GTACAGATAC	TCTACTACAC	TCAGCTCTCT	ATGTGCCAAG	5220
	TTTTTCTTTA	AGCAATGAGA	AATGCTCAT	GTCTTCTATC	TTCTCAAATC	ATCAGAGGCC	5280
	AAAGAAAAC	ACTTTGGCTG	TGCTATTAAC	TTGACACAGT	CAATGAAGTG	AAGAAAATTA	5340
35	GAGTAGTTAT	GTGATTATTT	CAGCTCTTGA	CCTGTCCCTC	GGGCTGCTCT	CTGAGTCTGA	5400
	ATCTCCCAAA	GAGAGAAACC	AATTTCTAAG	AGGACTGGAT	TGCAGAGAC	TGGGGGACAA	5460
	CATTGTATCC	AGATCTTAA	ATGTTATATT	GATAACCATG	CTCAGCAATG	AGCTATTAGA	5520
	TTCATTTTGG	GAAATCTCCA	TAAITTCAT	TTGTAACCTT	TGTTAAGACC	TGCTACATT	5580
40	GTATATGTTG	TGTGACTTGA	GTAATGTTAT	CAACGTTTTT	GTAATATTTT	ACTATGTTTT	5640
	TCTATTAGCT	AAATTCACAC	AATTTGTGAC	TTTAATAAAA	TGTTCTAAAC	ATTGC	

Seq ID NO: 178 Protein sequence:
Protein Accession #: NP_001058.1

45	1	11	21	31	41	51	
	MEVSPLQFVN	ENMQVNKIKK	NEDAKKRLSV	ERIVQKKTQL	EHILLRPDTY	IGSVELVTQQ	60
	MMVYDEDVGI	NYREVTVPFG	LYKIFDEILV	NAADNKQRDP	KMSCIRVTID	PENNLISINW	120
	NGKGIPVVEE	KVERMYVPAL	IFGQLLTSSN	YDDDEKQVTG	GRNGYGAKLC	NIFSTKFTVE	180
50	TASREYKMEF	KTRLMYDMGR	AGEMELKPFN	GEDYTCTITFQ	PDLSKFKMQS	LKDIDIVLMV	240
	RRAYDIAGST	KDVVFLNGN	KLFVKGFRSY	VDMLYKDKLD	ETGNSLKVLIH	EQVNRHWEVC	300
	LTMSEKGFQQ	ISFVSIATS	KGGRHVDYVA	DQIVTKLVDV	VKKKNKGGVA	VKAHQVKNHM	360
	WIFVNALIEN	PTFDSQTKEN	MTLQPKSFGS	TQQLSEKPIK	AAIGCGIVES	ILNWKVKFAQ	420
	VQLNKKCSAV	KHNRIKGIPIK	LDDANDAGGR	NSTECTLILT	EGDSAKTLAV	SGLGVVGRDK	480
55	YGVFPRLGKI	LNVRASHRQI	IMENAEINNI	IKIVGLQYKK	NYEDEDLSKT	LRYGKIMIMT	540
	DQDQDQSHIK	GLLNFNIHEN	WPSLLRHRFL	EEFITPIVKV	SKNKQEMAFY	SLPEFEENKS	600
	STPMHKWKV	KYKGLGTST	SKEAKEVFPD	MEKRIQFKY	SGPEDDAIS	LAFSKKQIDD	660
	RKEWLTNFM	DRRQRKLGL	PEDYLYGQTT	TYLTYNDFIN	KELILFNSND	NERSIPSMVD	720
60	GLKPGQKRVL	PTCFKRNDR	EVKVAQLAGS	VAEMSSYHHG	EMSLMNTIIN	LAQNPVGSNN	780
	LNLLQPIGQF	GTRLHGKDS	ASPRYIFTML	SSLARLLFPP	KDDHTLKLFLY	DDNQVPEPEW	840
	YIPIIPMVL	NGAEGIGTGW	SKKIPNFDVR	EIVNNIRRLM	DGEELPLMLP	SYKNFKGTIE	900
	ELAPNQYVIS	GEVAILNSTT	IEISELPVRT	WTQYKEQVL	EPMLNGTEKT	PPLITDYREY	960
	HTDTTVKFVV	KMTTEKLAEA	ERVGLHKVFK	LQTSLTQNSM	VLFDHVGLCK	KYDVTLDILR	1020
65	DFPELRKLY	GLRKEWLLGM	LGAESAKLNV	QARFILEKID	GKIIIEENPK	KELIKVLQIR	1080
	GYDSDPVKAW	KEAQKVPDE	EENEESDNEK	ETEKSDSVTD	SGPTFNLYLLD	MPLWYLTKKE	1140
	KDELCLRLNE	KEQELDTLKR	KSPSLWKED	LATFIEELEA	VEAKEKQDEQ	VGLPGKGGKA	1200
	KGKKTQMAEV	LPSPRQRVI	PRITIEKAE	AEKKNKKRIK	NENTEGSPQE	DGVELEGKQ	1260
	RLEKKQKREF	GTRTKKQTTL	AFKPIKGGKK	RNPWPDSESD	RSDSESNFDV	PPRETEPRRA	1320
70	ATKTKFTMDL	DSDEDFSDPD	EKTDDDFVP	SDASPPKTKT	SPKLSNKLK	PQKSVVSDLE	1380
	ADDVKGSVPL	SSSPPATHFP	DETEITNPVP	KKNVTVKKTA	AKSQSSTSTT	GAKKRAAPKG	1440
	TKRDPALNSG	VSQKPDPAKT	KNRRKRKPST	SDDSDSNFEK	IVSKAVTSKK	SKGESDDPHM	1500
	DFDSAVAPRA	KSVRAKKPIK	YLEESDEDDL	P			

Seq ID NO: 179 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 148-7095

80	1	11	21	31	41	51	
	CACACATACG	CACGACGAT	CTCACTTOSA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAG	ATTTCTCTTG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGGCGAGGGG	CCGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCTT	CGCTTGCAAT	180
	CAGCTCTCTT	GTGTTTGGCC	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
85	TCTGTGGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGGAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420

	AACACATTC	TTTCATACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTGAGGGAG	GAGTTTTCAGA	AATGGTGT	AAAGCAAGCA	AGATAACTTT	TCAGTGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
5	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTGTAGGA	AGCAGTCAAA	660
	GGAAAAAGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTGGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATCTCTGT	GAACCTTCTG	CCAACTCAA	CTGACAAATA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
10	ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTTTGTG	AAGTCTTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAAACAAT	TTGAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTT	CTCATACT	GGAAAGGAAG	AGATTTCATGA	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTACAA	1140
	TGGGAAAGAC	CTCGAGTCTG	TTATGATACC	ATGATTGAGA	AGTTTGAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
15	GGTGCTATTG	TCATAAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
20	ACGAAATCA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCTCTTTT	AAATGATGGC	TCTAAAACCTG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTGCG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
25	AGTTTATGCA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTCTCG	CTATCCCAAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAGAACCCG	AGCAATAATC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAAAG	TCACATAAAG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGTTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
30	AGCTTTCTCC	AGCTTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCAAG	2220
	TCCTTTTCTG	CAGGCCCACT	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTGCTTCA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTGCGT	CTCCACGGTC	AACTGGTAT	ACTGCGAGAC	AACCCAAACG	2400
35	GTATACAATG	GTGAGACACC	TCTTCAACCT	TCCTACAGTA	GTGAAGTCTT	TCCTTAGTGC	2460
	ACCCCTTTGT	TGCTTGACAA	TCAGATCCTC	AACACTACCC	CTGCTGCTTC	AAGTAGTGAT	2520
	TCGGCCTTGC	ATGCTACGCC	TGTATTTCCT	AGTGTGATG	TGTCAATTGA	ATCCATCCTG	2580
	TCCTCTCATG	ATGGTGCAAC	TTTGCTTCCA	TTTCTCTG	CTTCTTCAG	TAGTGAATTG	2640
	TTTCCGCCATC	TGCATACAGT	TTCTCAATC	CTTCCACAAG	TTACTTCAGC	TACCCAGAGT	2700
40	GATAAGGTGC	CCCTTGATGC	TTCTCTGCCA	GTGGCTGGGG	GTGATTGCT	ATTAGAGCCC	2760
	AGCCTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GAOCCTGGAA	2820
	TTTGTGATG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACCAACC	2880
	AGCAGTAGTG	CCATGATGCA	TGCACGTTCT	TCAGGGCCTG	AACCTTCTTA	TGCTTGTGCT	2940
	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTTCCTTACA	GTTCGTCAAT	ACCTGTGCAT	3000
	GATTTCTGTG	GTGTAACTTA	TCAGGGTTCC	TTATTTAGCG	GGCCTAGCCA	TATACCAATA	3060
45	CCTAAGTCTT	CGTTAAATAA	CCCAACTGCA	TCATTACTCT	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAAACAT	TTCTTCACTT	GTTTCTGTAG	CTGAATTTAC	ATATACAACA	3240
	TCTGTGTTTG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGGAATGAG	3300
	ACTGAAGTGC	AAATTTCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
50	CCCAACATAT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAGAAGAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAGGGGCAT	GTTTCCAGGG	TCCCTTGCTC	ATACCAACCAC	TAAAGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAT	AACITTTTCA	TTCAACCTAC	ACATACGTGC	3540
	TCTCAAGCAT	CTGGTGACAC	TTCCGTTAAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
	TCCTCTGACC	CTGCTTCTAG	TGAATGTTA	TCTCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
55	ACCTCAGCTT	CTTTTAGTAT	TGAAGTATTG	CTACAACCTT	CCTTTCAAGG	TTCTGATGTT	3720
	GACACCTTGC	TAAATACCTG	TCTTCCAGCT	GTGCCAGTG	ATCCAATATT	GGTGTAAACC	3780
	CCCAAGTTG	ATAAAAATTAG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
	AGTGAAAACA	TGCTGCACATC	TACATCTGTA	CCAGTTTGTG	ATGTTGCGCC	TACTTCTCAT	3900
	ATGCATCTCG	CTTCACTTCA	AGGTTTGACC	ATTTCTCATG	CAAGTGAGAA	ATATGAACCA	3960
60	GTTTGTGTTA	AAAGTGAAAG	TTCCCAACCA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAA	CAGGCCCATC	CCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATGTATGAA	CCATTAAATA	CACTAATAAA	TAAAGCTTATA	4140
	CATTCCGATG	AAATTTTAA	CTCCACCAAA	AGTTCTGTTA	CTGGTAAGGT	ATTTGCTGGT	4200
	ATTCCAAACG	TTGCTTCTGA	TACATTGTTA	TCTACTGATC	ATTCGTCTCC	TATAGGAAAT	4260
65	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCGTGTA	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACITCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAGGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
	AGAGGTAGTG	ATGGCTTATC	CATTATAAG	TGTATGTCAT	GCTCATCCTA	TAGAGAATCA	4500
	CAGGAAAAGG	TAAATGAATG	TTCAGACACC	CACGAAAACA	GTCTTATGGA	TCAGAAATAA	4560
70	CCAACTCTAT	ACTCACTATC	TGAGAAATCT	GAAGAAGATA	ATAGAGTCAC	AAGTGTATCC	4620
	TCAGACAGTC	AACTGGTAT	GGACAGAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACAATGATGG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCTGACAA	GTGATGAAGA	AAGTGGATCA	4800
75	GGGCAAGGTA	CCTCAGATAG	CCTTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTGTCA	4860
	GACACTAATG	AAAAAGATGC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AATAACTCCT	4920
	GGATTCCAC	AGTCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAGT	GTTCCACGTT	4980
	TCAGAGGCGA	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	5040
	GAATCCGAGA	AGATGGCACT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	5100
80	CTAGTGGTTC	TTTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	5220
	ATTTAGATG	ATGTCGAGC	AATTTCAATA	AAGCACTTTC	CAAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAAGT	5340
	CAGAGCTGTA	CTGTTGACTT	AGGTAATTACA	GCAGACAGCT	CCAACCAACC	AGACAACCAAG	5400
	CACAGAAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460
85	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	5520
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTC	5580
	TGGAGAATGA	TATGGGAACA	TAAATGTGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	5640

	AAAGGAAGGA	AAAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	5700
	TTTCTGGTCA	CTCAGAAGAG	TGTGCAAGTG	CTTGCCCTATT	ATACTGTGAG	GAATTTTACT	5760
	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	5820
5	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	5880
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTGCTG	5940
	CACGTGAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	6000
	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAAAACAT	CCGTTTCAAA	6060
	AGAAATTTAT	TGTTACAAAC	TGAGGAGCAA	TATGTCTTCA	TTCATGATAC	ACTGGTTGAG	6120
10	GCCTACTCTA	GTAAGAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCATGCTTA	TGTTAATGCA	6180
	CTCTCATTC	CTGGACCAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTTCA	GCTCCTGAGC	6240
	CAGTCAAAAT	TACAGCAGAG	TGACTATTCT	GCAGCCCTTA	AGCAATGCA	CAGGAAAAAG	6300
	AATCGAATCT	CTTCTATCAT	CCCTGTGGAA	AGATCAAGGG	TTGGCATTTC	ATCCCTGAGT	6360
	GGAGAAGGCA	CAGACTACAT	CAATGCCTCC	TATATCATGG	GCTATTACCA	GAGCAATGAA	6420
	TTCTACTCTA	CTGACACCCC	TCTCCTTCAT	ACCATCAAGG	ATTTCTGGAG	GATGATATGG	6480
15	GACCATTAAT	CCCAACTGGT	GGTTATGATT	CCTGATGGCC	AAAACATGGC	AGAAGATGAA	6540
	TTTGTTTACT	GGCCAAATAA	AGATGAGCCT	ATAAATTGTG	AGAGCTTTAA	GGTCACTCTT	6600
	ATGGCTGAAG	AACACAAATG	TCTATCTAAT	GAGGAAAAAC	TTATAATTCA	GGACTTTTATC	6660
	TTAGAAGCTA	CACAGGATGA	TTATGTACTT	GAAGTGAGGC	ACTTTTCAGT	TCCTAAATGG	6720
20	CCAAATCTAG	ATAGCCCATC	TAGTAAAACT	TTTGAACCTA	TAAGTGTATT	AAAAGAAGAA	6780
	GCTGCCAATA	GSGATGGGCC	TATGATTGTT	CATGATGAGC	ATGGAGGAGT	GACGGCAGGA	6840
	ACTTTCTGTG	CTCTGACAAC	CCTTATGCAC	CAACTAGAAA	AAGAAAATTC	CGTGGATGTT	6900
	TACCAGTAGT	CCAGATGATG	CAATCTGATG	AGGCCAGGAG	TCTTTGCTGA	CATTGAGCAG	6960
	TATCAGTTTC	TCTACAAAGT	GATCCTCAGC	CTTGTGAGCA	CAAGGCAGGA	AGAGAATCCA	7020
25	TCACCTCTC	TGGACAGTAA	TGGTGACGCA	TTGCTGATG	GAAATATAGC	TGAGAGCTTA	7080
	GAGTCTTTAG	TTTAAACACAG	AAAGGGGTGG	GGGGACTCAC	ATCTGAGCAT	TGTTTCTCTC	7140
	TTCTTAAAT	TAGGCAGGAA	AATCAGTCTA	GTTCTGTTAT	CTGTTGATT	CCCATCACCT	7200
	GACAGTAAGT	TTTATGACAT	AGGATTCTGC	CGCCAAATTT	ATATCATTAA	CAATGTGTGC	7260
	CTTTTGTCAA	GACTTGTAAT	TTACTTATTA	TGTTTGAAGT	AAAATGATTG	AATTTTACAG	7320
30	TATTTCTAAG	AATGGAAATG	TGGTATTTT	TTCTGTATTG	ATTTTAAACG	AAAATTTCAA	7380
	TTTATAGAGG	TTAGGAATTC	CAAACTACAG	AAAATGTTTG	TTTTTAGTGT	CAAAATTTTA	7440
	GCTGTATTGG	TAGCAATTTAT	CAGGTTTGCT	AGAAATATAA	CTTTTAATAC	AGTAGCCTGT	7500
	AAATAAAACA	CTCTTCCATA	TGATATTCAA	CATTTTACAA	CTGCAGTATT	CACCTAAAGT	7560
	AGAAATAATC	TGTTACTTAT	TGTAATAACT	GCCCTAGTGT	CTCCATGGAC	CAAAATTTATA	7620
35	TTTATAATGG	TAGATTTTTA	TATTTTACTA	CTGAGTCAAG	TTTTCTAGTT	CTGTGTAATT	7680
	GTTTAGTTTA	ATGACGTAGT	TCATTAGCTG	GTCCTACTCT	ACCAGTTTTT	TGACATTGTA	7740
	TTGTGTTACC	TAAAGTCATTA	ACTTTGTTTC	AGCATGTAAT	TTTAACTTTT	GTGAAAAATA	7800
	GAAATACCTT	CATTTTGAAG	GAAATTTTTA	TGAGAATAAC	ACCTTACCAA	ACATTTGTCA	7860
40	AATGGTTTTT	ATCCAAGGAA	TTGCAAAAAT	AAATATAAAT	ATTGCCATTA	AAAAAAAATA	7920
	AAAAAAAATA	AAAAAAAATA	AAAA				

Seq ID NO: 180 Protein sequence:
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
45	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGNSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQNVNVLKKL	KFQGWDKTSL	ENTFIHNTGK	TVEINLNDY	RVSGGVSEMV	120
	FKASKITFW	GKCNMSDGS	EHSLEGGKFP	LEMQIYCFDA	DRFSSFEAV	KGKGLRLALS	180
	ILFEVGTSEN	LDPKAIDGV	BSVSRPGKQA	ALDPFILLNL	LPNSTDKYVI	YNGSLTSPPC	240
50	TDVDMIVFK	TDVSISESQ	AVFCEVLTMQ	QSGYVLMMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGKEETHEAV	CSSEPNVQA	DPENYTSLLV	TWERPRVVDY	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNNLPL	NMSYVLQIVA	ICTNGLYGKY	SDQLIVDMPT	DNPELDLPPE	420
	LIGTEELIKE	EEEGKIDIEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTTHYNRI	GTRYNEAKTN	480
	RSPTRGSEFS	GKGDVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPT	VEGTSASLND	540
55	GSKTVLRSPH	MNLSGTAEBSL	MTVSITEYEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIPS	SENPEITITD	VLIPEASRNA	SEDSTSSGSE	ESLKDPSEMG	NWPFSSSTDI	660
	TAQPDVGSGR	ESFLQNTYTE	IRVDESEKTT	KSFSAGPVMS	QGPSVTDLEM	PHYSTFAFP	720
	TEVTPHAFPT	SSRQDLVST	VNVVYSQTTQ	PVYNGETPLQ	PSYSSEVPPL	VTPLLLDNQI	780
	LNTTPAASSS	DSALHATPVF	PSVDVSFESI	LSSYDGAPLL	PFSSASFSE	LFRHLHTVSQ	840
60	ILPQVTSATB	SDKVP LHASL	PVAGDLLLLL	PSLAQYSDVL	STTHAASETL	EFGSESGVLY	900
	KTLMPQSVEP	PSSDAMMHAR	SSGPEPSYAL	SDNEGSQHIF	TVSYSSAIPV	HDSVGVTYQG	960
	SLFSGPSHIP	IPKSSLTITPT	ASLQPTHAL	SQDGEWSGAS	SDSEFLLED	DGLTALNISS	1020
	PVSVAEFTYT	TSVPQDDNKA	LSKSEIIYGN	ETELQIPSPN	EMVVPSESTV	MFNMYDNVVK	1080
	LNASLOETSV	SISSTKGMFP	GSLAHTTTKV	FDHEISQVPE	NNPVSQPTHT	VSQASGDTSL	1140
65	KPVLNANSEP	ASSDPASSEM	LSPSTQLLFY	ETSASFSTEV	LLQPSFQASD	VTLLKTVLPL	1200
	AVPSDPIVLE	TPKVDKISST	MLHLIVNSA	SSEMLHSTS	VPVPDVSPTS	HMSASLQGL	1260
	TISYASEKYE	PVLLKSESSH	QVPSLYSND	ELFQNTANLEI	NQAHPPKGRH	VFATPVLISID	1320
	EPLMTLNLKL	IHSDEILTST	KSSVTGKVFA	GIPTVASDTP	VSTDHSVPIG	NGHVAITAVS	1380
	PHRDGSVTST	KLFPFSKATS	ELSHSAKSDA	GLVGGGEDGD	TDDGDDDDDD	DRGSDGLSIH	1440
70	KCMSCSSYRE	SQEKVMNDSD	THENSLMDQN	NPISYSLSEN	SEEDNRVTIS	SSDSQTGMDR	1500
	SPGKSPSANG	LSQKHNDGKE	ENDIQTSAL	LPLSPESKAW	AVLTSDEESG	EQGTSDSLNL	1560
	ENETSDPSFR	ADTNEKDADG	ILAAQDSEIT	PGFPQSPTSS	VTSENSEVFH	VSEAEASNSS	1620
	HESRIGLAEG	LESEKKAVIP	LVIVSALTPI	CLVVLVGILI	YWRKCFQTAH	FYLEDSTSPR	1680
	VISTPPTPIF	PISDDVGAIP	IKHFPKHVAD	LHASSGFTSE	FETLKEFYQE	VQSCTVDLGI	1740
75	TADSSNHPDN	KHKRYINIV	AYDHSRVKLA	QLAEKDGKLT	DYINANYVDG	YNRPKAYIAA	1800
	QGPKLSTAE	FWRMWHEHNV	EVIVMITNLV	EKGRKCDQY	WPADGSEBYG	NFLVTQKSVQ	1860
	VLAYYTVRNP	LTNTKIKKG	SQKGRPSGRV	VTQYHYTQWP	DMGVPEYSLP	VLTVPVRKAA	1920
	AKRHAQVPPV	VHCSAGVGR	GTIVVLSML	QQIQHEGTVN	IFGFLKHIRS	QRNYLVQTEB	1980
	QYVFIHDTLV	EALISKETE	LDSHIHAYVN	ALLIPGPAGK	TKLEKQFQLL	QSNIQSQSDY	2040
80	SAALKQCNRE	KNRTSSIIIP	ERSRVGISSL	SEEGTDYINA	SYIMGYYQSN	EFITQHPILL	2100
	ETIKDFWRMI	WDHNAQLVVM	IPDQGNMAED	EFVYWNPKDE	PINCESPKVT	LMAEHLKCLS	2160
	NEEKLIQDF	ILEATQDDYV	LEVRFQCPK	WPNPDSPIK	TFELISVIKE	BAANRDGPMI	2220
	VHDEHGGVTA	GTFCALTTLM	HQLEKENSVD	VYQVAKMINL	MRPGVFADIE	QYQFLYKVIL	2280
85	SLVSTQREEN	PSTSLDSNGA	ALPDGNIAES	LESIV			

Seq ID NO: 181 DNA sequence
Nucleic Acid Accession #: Eos sequence

	1	11	21	31	41	51	
5	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCTCTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCGCGA	120
	CGCGCAGGGG	CGCGCAGCCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
10	CTTGTGGAAG	AGATTGGCTG	GTCTATACA	GGAGCACTGA	ATCRAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTTGAA	420
	AACACATTCA	TTTATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTGAGCGGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCATCTGGGA	540
15	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TGCGGACAGA	AGAAAATTTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATCTGTTC	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTCAC	840
20	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTGTGTG	AAGTCTTTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAAAT	TTCCAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAG	AGGTGTGTTT	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
	AGTTTCAAGC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTGTACA	1140
25	TGGGAAAGAC	CTGAGTCTGT	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTTGTACCAAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTC	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACATAAT	GCTTATATGG	AAAATACAGC	GACCAACTGA	TGTGCGACAT	GCCTACTGAT	1380
	AATCTCGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACTG	AAGAAATAAT	CAAGGAGGAG	1440
30	GAAGAGGAA	AGACATTTGA	AGAAGGCGCT	ATTGTGAATC	CTGTAGAGAG	CAGTGTCTACA	1500
	AACCAAAATC	GGAAAAGGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TGCGATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTAGTG	TTCCCAATAC	ATCTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAGG	ATATTCTCTT	GACTTCTCAG	ACTGTGACTG	AAGTCCACCC	TCACACTGTG	1740
35	GAAGGTACT	CAGCTCTCTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCCATATG	1800
	AACCTGTGCG	GGAGTCGAGA	ATCCTTAAAT	ACAGTTTCTA	TACAGAAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAGAACCCG	AGCAATAAAC	ATATGATGTC	CTTATACCCG	AATCTGCTAG	AAATGCTTCC	2040
40	GAAGATTCAA	CTTCACTCAG	TTCAAGAGAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAAATTA	CATGAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCAAG	2220
	TCCTTTCTG	CAGGCCAGAT	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTGCTCTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
45	TCCAGACACA	AGGATTGTTG	CTCCACGGTC	AACGTGGTAT	ACTGCGACAG	AACCCAAACG	2400
	GTATACAATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGCTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAGAGC	AGTTATACCC	CTTGTGATCG	TGTGAGCCCT	GACTTTTATC	2520
	TGTCTAGTGG	TTCTTGTGGG	TATCTCATC	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
	TTTTACTTAG	AGGACAGTAC	ATCCCTTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTT	2640
50	CCAATTCTAG	ATGATGTCCG	AGCAATTTCCA	ATAAGCACT	TTCCAAAGCA	TGTTGAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTGAAGAA	TTTGAGACAC	TGAAAGAGTT	TTACAGGAA	2760
	GTGCAGAGCT	GTACTGTTGA	CTTAGGTATT	ACAGCAGACA	GCTCCAACCA	CCGAGACAAC	2820
	AAGCAACAAG	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAAAGCTAGCA	2880
	CAGCTTGCTG	AAAAGGATGG	CAAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
55	TACAACAGAC	CAAAAGCTTA	TATTTGCTGC	CAAGGCCAC	TGAATCCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAGTTATTG	TCATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAAATG	TGATCAGTAC	TGGCCTCGCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACCTTCTGG	TCACTCAGAA	GAGTGTGCAA	GTGCTTGCC	ATTATACCTG	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAAAT	AAAAAGGGGC	TCCACAGAA	GAAGACCCAG	TGGACGTGTG	3240
60	GTCCACAGAT	ATCACTACAC	GCAGTGGCCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGTGAGAAA	GGCAGCCTAT	GCCAAAGGCC	ATGCAGTGGG	GCCTGTTGTC	3360
	GTCCACTGCA	GTGCTGGAGT	TGGGAAGACA	GGCACATATA	TGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTG	AACACGAAGG	AACGTCTAAC	ATATTGGCT	TCTTAAACAA	CATCCGTACA	3480
	CAAGAAATTT	ATTTGATACA	AACGTAGGAG	CAATATGCT	TCAATCATGA	TACACTGGTT	3540
65	GAGGCCATAC	TTAGTAAAGA	AACGTAGGAG	CTGACAGTGC	ATATTCTATG	CTATGTTAAT	3600
	GCACCTCTCA	TTCTGGAGCC	AGCAGGCAAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGACTAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAATCGAA	CTTCTCTAT	CATCCCTGTG	GAAAGATCAA	GGGTTGGCAT	TTTATCCCTG	3780
	AGTGGAGAGG	GCACAGACTA	CATCAATGCC	TCCTATATCA	TGGGCTATTA	CCAGAGCAAT	3840
70	GAATTCATCA	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTCTTG	GAGGATGATA	3900
	TGGGACCATTA	ATGCCCAACT	GGTGGTTATG	ATTCTCTGAT	GCCAAAACAT	GGCAGAAAGAT	3960
	GAATTTGTTT	ACTGGCCAAA	TAAAGATGAG	CCTATAAAAT	GTGAGAGCTT	TAAAGTCACT	4020
	CTTATGGCTG	AAGAACAACA	ATGCTATCT	AATGAGGAAA	AACCTTAAAT	TCAGSACTTT	4080
	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTTGAAGTGA	GGCACTTTCA	GTGCTCTAAA	4140
75	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAAGAA	4200
	GAAGCTGCCA	ATAGGGATGG	GCCTATGATT	GTTTATGATG	AGCATGGAGG	AGTGACGGCA	4260
	GGAACTTTCT	GTGCTCTGAC	AACCTTATG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCCAG	TAGCCAGAGT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTTG	TGACATTGAG	4380
	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTTG	GCACAAGGCA	GGAGAGAAAT	4440
80	CCATCCACCT	CTCTGAGACG	TAAATGGTGA	GCATTGCTCT	ATGGAAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAAC	CAGAAAGGGG	TGGGGGAGCT	CACATCTGAG	CATTGTTTTT	4560
	CTCTTCTTAA	AAATTAGGCG	GAAATTCAGT	CTAGTCTCTG	TATCTGTTGA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTTTATGA	CATAGGATTC	TGCCGCCAAA	TTTATATCAT	TAACAATGTG	4680
	TGCCCTTTTG	CAAGACTTGT	AAATTTACTTA	TTATGTTTGA	ACTAAAATGA	TGAAATTTTA	4740
85	CAGTATTCTT	AGAATAGGAA	TTGTGGTATT	TTTTTCTGTA	TGATTTTAA	CAGAAAATTT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAAACTA	CAGAAAATGT	TTGTTTTTAA	TGTCAAAATTT	4860
	TTAGCTGTAT	TTGTAGCAAT	TATCAGGTTT	GCTAGAAATA	TAACTTTTAA	TACAGTAGCC	4920
	TGTAATAAAA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCAGT	ATTCACCTAA	4980

5 AGTAGAATA ATCTGTACT TATTGTAAAT ACTGCCCTAG TGCTCCCATG GACCAAAATT 5040
 ATATTATATA TTGTAGATT TTATATTTTA CTACTGAGTC AAGTTTCTA GTTCTGTGTA 5100
 ATTGTTTAGT TTAATGACGT AGTTCATTAG CTGGCTCTAC TCTACCAGTT TTCTGACATT 5160
 GTATTGTGTT ACCTAAGTCA TTAACCTTGT TTCAGCATGT AATTTTAACT TTGTGGGAAA 5220
 ATAGAAATAC CTTCAATTG AAAGAAGTTT TTATGAGAAT AACACCTTAC CAAACATTGT 5280
 TCAATGGTT TTTATCCAAG GAATTGCAAA AATAAATATA AATATTGCCA TTAATAAAAA 5340
 AAAAAAAAAA AAAAAAAAAA AAAAAA

10 Seq ID NO: 182 Protein sequence:
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 15 MRILKRLFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALEQKNWG KKYPTCNSPK 60
 QSPINIDEDL TQVNVNLLKL KFGQWDRKSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120
 FKASKITFFH GKCMSSSDGS EHSLEGQKFP LEMQIYCFDA DRPSSPEEAV KKGKGLRLALS 180
 ILFEVGTSEN LDFKAIIDGV ESVSFPGKQA ALDPFILLNL LPNSTDKYIY YNGSLTSPPC 240
 TDTVDWIVFK DTVSISESL AVFCEVLTMO QSGYVMLMDY LQNNFREQQY KFSRQVFSSY 300
 20 TGKEEIEHAV CSSEPEVQA DPENYTSLLV TWERPRVVD TMIEKFAVLY QQLDGEDQTK 360
 HEFLTDGYQD LGATLNNLLP NMSYVLQIVA ICTNGLYKGY SDQLIVDMPT DNPFLDLFPE 420
 LIGTEELIKE EEEGKLIBEG AIVNFGKDSA TNQIRKKEPQ ISTTTHYNI RTKYNEAKTN 480
 RSPTRGSEFS KGGVDPNTSL NSTSQPVTKL ATEKDISLTS QVTLELPHT VEGTSASLND 540
 GSKTVLRSPH MNLGTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600
 25 ENISQGYIFS SENPETITYD VLIPEASRNA SEDSTSSGSE ESLKDPMEG NVWPPSSIDI 660
 TAQPDVGSGR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGSPVTDLEM PHYSTFAYFP 720
 TEVTPHAFTE SSRQDLVST VNVVYSQTTQ PYNNAEASNS SHESRIGLAE GLESEKKAIV 780
 PLVIVSALT ICLVVLVGIL IYWRKCPQTA HPLYEDSTSP RVISTPPTPI FPISDDVGAI 840
 30 PIKHPKHA DLHASSGFTE EPETLKEPYQ EVQSCIVDLG ITADSSNHPD NKEKNRYINI 900
 VAYDHSRVLK AQLAEKDKGL TDYINANYVD GYNRPKAYIA AQGPLKSTAE DFWPMIWEHN 960
 VEVIVMITNL VEKGRKCDQ YWPADGSEEV GNFLVTQKSV QVLAYTYVRN FTLRNTKIKK 1020
 GSQGRPSGR VVTQYHYTQW PDMGVPEYSL PVLTFVRKAA YAKRHAVGPV VVHCAGVGR 1080
 TGTIVLDSM LQIQHEGTV NIFGFLKHIR SQRYNLVQTE EQYVPHDTL VEAILSKETE 1140
 35 VLDSHIHAYV NALLIPGPAG KTKLEKQFQL LSQSNIQQSD YSAALKQCNR EKNRTSSIIP 1200
 VERSRVGISS LSGGTDYIN ASYIMGYYS NEFIITQHPL LHTIKDFWRM IWDHNAQLVV 1260
 MIDPGNMAB DEFVYWNPKD EPINCESFKV TLMAEHEKCL SNEEKLIQD FILEATQDDY 1320
 VLEVRHPCPE KWPNDPSPI KTFELISVIK EEAANRDGPM IVDEHGGVT AGTFCALITL 1380
 MHQLEKENS SVYQVAKMIN LMRPGVFADI EQYQFLYKVI LSLVSTRQEE NPSTSLDSNG 1440
 40 AALPDGNIAE SLESIV

Seq ID NO: 183 DNA sequence
 Nucleic Acid Accession #: EOS sequence
 Coding sequence: 148-4494

45 1 11 21 31 41 51
 CACACATAG CACGACGAT CTCACTTGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAAC ATTTCCTTCG CTCCCTCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
 50 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCTTAA AGCGTTTCTT CGCTTGCATT 180
 CAGCTCCTCT GTGTTTGGCG CTGTAATGGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTGTGTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG 300
 AAATATCCAA CATGTAATAG CCAAAACAAA TCTCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCAITGGAA 420
 AACACATTTA TTTAATACAC TGGGAAAACA GTGGAAATTA ATCTCACTAA TGACTACCGT 480
 55 GTCAGCGGAG GAGTTTTCAGA AATGGTGTIT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540
 AAATGCAATG TGTCATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600
 GAGATGCAAA TCTACTGCTT TGATGCAGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660
 GGAAGAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAAATTG 720
 60 GATTTCGAAG CGATTATTTA TGGAGTCGAA AGTGTATAGT GTTTTGGGAA GCAGGCTGCT 780
 TTAGATCAAG TCATCTGTTT GAACCTTCTG CCAACTCTCA CTGACAAGTA TTACTTTTAC 840
 AATGGCTCAT TGACATCTCC TCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTGTG AGTTTCTTAC AATGCAACAA 960
 TCTGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTCGAGAGCA ACAGTACAAG 1020
 65 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTCTAGA AGCAGTTTGT 1080
 AGTTCAGAAC CAGAAAAATG TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA 1140
 TGGGAAGAGC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCACT TTTGTACCAG 1200
 CAGTTGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATTG TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 70 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGAGCAT GCCTACTGAT 1380
 AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAACAG AAGAAATAAT CAAGGAGGAG 1440
 GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500
 AACCAATCA GGAAGAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCCCATAGGG 1560
 75 ACAGAAATACA ATGAAGCCAA GACTAACCGA TCCCAACAAA GAGGAAGTGA ATTCTCTGGA 1620
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
 ACAGAAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGGTACTT CAGCTCTTTT AAATGATGCG TCTAAAACCTG TTCTTAGATC TCCACATATG 1800
 AACTTGTGCG GGACTGCAGA ATCCTTAAT ACAGTTTCTA TAACAGAAAT TGAGGAGGAG 1860
 80 AGTTTATGCA CCAAGTTTCA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920
 GCAACTTCTG ATATCCCAAT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980
 GAAAACCCAG ACAGCAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040
 GAAGATTCAA CTTATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAT 2100
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160
 AGCTTCTCC AGACTTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220
 85 TCTTTTCTG CAGGCCAGT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAATAGCCA 2280
 CATTATCTA CTTTGCCTA CTCCCAACT GAGGTAACAC CTCATGCTTT TACCCATCC 2340
 TCCAGACAAC AGGATTGGT CTCCAGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG 2400
 GTATACAATG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTTG 2460

	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
5	ATTTACAGAT	ATGTCGGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGSTTTAC	TGAAGAATTT	GAGGAAGTGC	AGAGCTGTAC	TGTTGACTTA	2760
	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAAACAGC	ACAAGAATCG	ATACATAAAT	2820
	ATCGTTGCC	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCAAA	2880
	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
10	GCTGCCCAAG	GCCCCTGAA	ATCCACAGCT	GAGGATTCTT	GGAGAATGAT	ATGGGAACAT	3000
	AATGTGGAAG	TTATTGTCAT	GATAACAAAC	CTCGTGGAGA	AAGGAAGGAG	AAATGTGAT	3060
	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAC	TTCTGGTCAC	TCAGAAGAGT	3120
	GTGCAAGTGC	TTGCCATTAT	TACTGTGAGG	AATTTTACTC	TAAGAAACAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGG	CGTGTGGTCA	CACAGTATCA	CTACACCGAG	3240
	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGSCA	3300
15	GCCATATGCC	AGGCCCATGC	AGTGGGCTCT	GTGTCGTGCC	ACTGCAGTGC	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAAGT	3420
	GTCAACATAT	TTGGCTTCTT	AAAAACATC	CGTTCAACAA	GAATTTATTT	GGTACAAACT	3480
	GAGGAGCAAT	ATGCTTTCAT	TCATGATACA	CTGGTTGAGG	CCATACCTAG	TAAGAAACT	3540
20	GAGGTGCTGG	ACAGTCATAT	TCATGCTTAT	GTTAATGCAC	TCCTCATTCC	TGGACCAAGC	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCCAG	CTCCTGAGCC	AGTCAAAATAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAGAA	ATCGAACTTC	TTCTATCATC	3720
	CCTGTGGAAG	GATCAAGGAT	TGGCAATTCA	TCCCTGAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGCTCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCATTAC	CCAGCACCCCT	3840
25	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	ACCATAATGC	CCAATGGTGG	3900
	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
	GATGAGCCTA	TAAATTTGTA	GAGCTTTAAG	GTCACCTCTA	TGGCTGAAGA	ACACAAATGT	4020
	CTATCTAATG	AGGAAAACCT	TATAATTGAG	GACITTTATCT	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	CAATCCAGA	TAGCCCCATT	4140
30	AGTAAACCTT	TTGAACCTAT	AAGTGTATTA	AAAGAAGAAG	CTGCCAATAG	GGATGGGCTT	4200
	ATGATTTGTT	CTGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	CTTTCTGTGC	TCTGACAAAC	4260
	CTTATGACCC	AACATAGAAA	AGAAAATTC	GTGGATGTTT	AOCAGGTAGC	CAAGATGATC	4320
	AATCTGAATG	GCCAGGAGGT	CTTGTCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACCTCTCT	GGACAGTAAT	4440
35	GGTGACAGAT	TGCCTGATGG	AAATATAGCT	GAGAGCTTAG	AGTCTTTAGT	TTAACACAGA	4500
	AAGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTTCTCT	TCCTAAATTT	AGGCAGGAAA	4560
	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTC	CCATCACCTG	ACAGTAACTT	TCATGACATA	4620
	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTTGTC	TTTTTGCAAG	ACTTGTAAAT	4680
	TACTTATTAT	GTTTGAACCT	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
40	GGTATTTTTT	TCTGTATATG	TTTTAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTCC	4800
	AAACTACAGA	AAATGTTTGT	TTTTAGTGT	AAATTTTAT	CTGTATTGTT	AGCAATTATC	4860
	AGGTTTGCTA	GAAATATAAC	TTTTAATACA	GTAGCCTGTA	AATAAACAC	TCTTCCATAT	4920
	GATATTCAAC	ATTTACAAC	TGCAGTATTC	ACCTAAAGTA	GAAATAATCT	GTTACTTATT	4980
	GTAATACTG	CCCTAGTGT	TCCATGGACC	AAATTTATAT	TTATAATTGT	AGATTTTAT	5040
45	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATTG	TTTAGTTTAA	TGACGTAGTT	5100
	CATTAGCTGG	TCTTACTCTA	CCAGTTTCT	GACATTGTAT	TGTGTACCT	AAGTCATATA	5160
	CTTGTGTTCA	GCAATGAAT	TTAATCTTTG	TGGAAATAG	AAATACCTTC	ATTTTGAAG	5220
	AAGTTTATAT	GAGAAATACA	CCTTACCRAA	CATTGTTCAA	ATGGTTTTTA	TCCAAGGAAT	5280
	TGCAAAAAAT	AATATAAATA	TTGCCATTAA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA	5340
50	AAA						

Seq ID NO: 184 Protein sequence:
Protein Accession #: EOS sequence

	1	11	21	31	41	51	
55	MRILKRFAC	IQLLCVRLD	WANGYYRQR	KLVEEIGNSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDDEL	TQVNVNKLK	KFGQWDTSL	ENTFIHNTGK	TVEINLTNDY	RVSQGVSEMV	120
	FKASKITFW	GKCNMSDGS	EHSLEGQKFP	LEMQIYCFDA	DRPSSPREAV	KGKGLRLALS	180
60	ILFEVGTZEN	LDFKAIIDGV	ESVSRPGKQA	ALDPPILLNL	LPNSTDKYI	YNGSLTSPPC	240
	TDVDWIVFK	DTVSISESQ	AVFCEVLTMO	QSGVVMMDY	LQNNFREQQY	KPSRQVPSY	300
	TGREIEHEAV	CSSPENVOA	DPENYTSLLV	TWERPRVVD	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTQDYG	LQAILNLLP	NMSYVLQIVA	ICTNGLYGYK	SDQLIVDMPT	DNPELDLFE	420
	LIGTEELIKE	BEEGKIDIEG	AIVNPGRDSA	TNQLRKKEPQ	ISTTTTHYNI	GTYKNEAKTN	480
65	RSPTRGSEPS	GKGDVPTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPFPH	VEGTASLND	540
	GSKTVLRSPH	MNLSGTAESE	NTVSIITEYEE	ESLLTSPKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETITVD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NWVFPSTDI	660
	TAQPDVSGSR	ESFLQTYMYE	IRVDESEKTT	KSFSAGPVMS	QGPSVIDLEM	PHYSTPAYFP	720
	TEVTPHAFPT	SSRQDLVST	VNVVYSQTTQ	PVYNEASNSS	HESRIGLAEG	LESEKKAVIP	780
70	LVIVSALTFI	CLVVLVGLI	YWRKCFQTAH	FYLEDSTSPR	VISTPPTPIF	PISDDVGAIP	840
	IKFPPKHVAD	LHASSGFTEE	FEEVQSCTVD	LGITADSSNH	PDNKHKNRYI	NIVAYDHSRV	900
	KLAQLAEKDG	KLTYINANY	VDGYNRPKAY	IAAQGLPKST	AEDFWRMIWE	HNVFVIVMIT	960
	NLVEKGRKRC	DQYWPADGSE	EYGNFLVTQK	SVQVLAYYTV	RNFTLRNTKI	KKGSQKGRPS	1020
	GRVVTQYHYT	QWPMGVPEY	SLPVLTFVRK	AAYAKRHAVG	PVVVHCSAGV	GRTGTYIVLD	1080
75	SMLQIQIHG	TVNIFGFLKH	IRSQRNVLQV	TEEQYVFIHD	TLVEAILSKE	TEVLDSHIIA	1140
	YVNALLIPGF	AGKIFGLEKQ	QLLSQSNLQ	SDYSAALKQC	NREKRTSSI	IPVERSRVGI	1200
	SSLSGEGTDY	INASYIMGYV	QSNFPIITQH	PLLHTIKDFW	RMIWDHNAQL	VVMIPDQGNM	1260
	AEDEFVYWN	KDEPINCESF	KVTLMAREHK	CLSNEEKLII	QDFILEATQD	DYVLEVRHFQ	1320
	CPKWPNDPSP	ISKTFELISV	IKEEAANRDG	PMIVHDEHGG	VTAGTFPCALT	TMHQLKEKEN	1380
80	SVDVYQVAKM	INLMRPQVFA	DIEQYQFLYK	VILSLVSTRQ	EENPSTSLDS	NGAALPDGNI	1440
	AESLESLSV						

Seq ID NO: 185 DNA sequence
Nucleic Acid Accession #: EOS sequence
Coding sequence: 501-4514

85	1	11	21	31	41	51
----	---	----	----	----	----	----

	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGCGAGGGGG	CCGAGAGCCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
5	CAGCTCCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCCATATACA	GGAGCACTGA	ATCAAAAAAT	TGGGGAAGA	300
	AATATCCAACT	ATGTAATAGC	CCAAAACAACT	CTCCTATCAA	TATTTGATGA	GATCTTACAC	360
	AGTAAATGT	GAATCTTAAG	AAACTTAAAT	TTCAGGGTTG	GGATAAAACA	TCATTGGAAA	420
	ACACATTCAT	TCATAACACT	GGGAAAACAG	TGGAAATTAA	TCTCACTAAT	GACTACCGTG	480
10	TCAGCGGAGG	AGTTTCAGAA	ATGGTGTTTA	AAGCAAGCAA	GATAACTTTT	CACTGGGGAA	540
	AATGCAATAT	GTCACTGTAT	GGATCAGAGC	ATAGTTTAGA	AGGACAAAAA	TTTCCACTTG	600
	AGATGCAAT	CTACTGCTTT	GATCGGAGCC	GATTTTCAAG	TTTTGAGGAA	GCAGTCAAGG	660
	GAAGAGGGAA	GTTAAGAGAT	TTATCCATTT	TGTTTGAAGT	TGGGACAGAA	GAATAATTGG	720
	ATTTCAAAGC	GATTATGTAT	GGAGTCGAAA	GTGTTAGTCG	TTTTGGGAAG	CAGGCTGCTT	780
15	TAGATCCATT	CATACGTGTT	AACCTTCTGC	CAAACTCAAC	TGACAGTAT	TACATTTACA	840
	ATGGCTCATT	GACATCTCCT	CCCTGCACAG	ACACAGTTGA	CTGGATTGTT	TTTAAAGATA	900
	CAGTTAGCAT	CTCTGAAGCG	CAGTTGGCTG	TTTTTGTGTA	AGTTTCTACA	ATGCAACAAT	960
	CTGGTTATGT	CATGCTGATG	GACTACTTAC	AAAACAATTT	TGAGAGCAAA	CAGTACAAGT	1020
20	TCTCTAGACA	GGTGTTTTCC	TCATACACTG	GAAGGAAGA	GATTCAATGA	GCAGTTTGTG	1080
	GTTCAGAACCC	AGAAAATGTT	CAGGCTGACC	CAGAGAATTA	TACCAGCCTT	CTTGTTTACAT	1140
	GGGAAAGACC	TCCAGTCGTT	TATGATACCA	TGATTGAGAA	GTTTGCAGTT	TGTATCCAGC	1200
	AGTTGATGG	AGAGGACCAA	ACCAAGCATG	AATTTTGGAC	AGATGGCTAT	CAAGACTTGG	1260
	GTGCTATTCT	CAATAATTGG	CTACCCAAAT	TGAGTTATGT	TCTTCAGATA	GTAGCCATAT	1320
	GCATAATGG	CTTATATGGA	AAATACAGCG	ACCAACTGAT	TGTGACATG	CCTACTGATA	1380
25	ATCCTGAAC	TGATCTTTTC	CCTGAATTAA	TGGGAAGTGA	AGAAATATC	AAGGAGGAGG	1440
	AAGAGGGAAA	AGCATTGTAA	GAAGGCGCTA	TGTGAAATCC	TGGTAGAGAC	AGTGCTACAA	1500
	ACCAATACG	GAATAAGGAA	CCCCAGATTT	CTACCAACAAC	ACACTACAAT	CGCATAGGGA	1560
	CGAAATACAA	TGAAGCCAA	ACTAACCGAT	CCCCAACAA	AGGAAGTGAA	TTCTCTGGAA	1620
	AGGGTGATGT	TCCCAATACA	TCTTTAAAT	CCACTTCCCA	ACCACTCACT	AAATTAGCCA	1680
30	CAGAAAAAGA	TATTTCTTGG	ACTTCTCAGA	CTGTGACTGA	ACTGCCACCT	CACACTGTGG	1740
	AAGGTACTTC	AGCCTCTTTA	AATGATGGCT	CTAAACTGT	TCTTAGATCT	CCACATATGA	1800
	ACTTGTGCGG	GACTGTCAGAA	TCCITAAATA	CAGTTTCTAT	AACAGAAAT	GAGGAGGAGA	1860
	GTTTATTGAC	CAGTTTCAAG	CTTGATCTG	GAGCTGAAGA	TTCTTCAGGC	TCCAGTCCCG	1920
35	CAACTTCTGC	TATCCCATTC	ATCTCTGAGA	ACATATCCCA	AGGGTATATA	TTTTCTCTCG	1980
	AAAACCCAGA	GACAAATAACA	TATGATGTCC	TTATACCAGA	ATCTGCTAGA	AATGCTTCCG	2040
	AAGATTCAAC	TTATCTAGGT	TCAGAGAAT	CACTAAAGGA	TCCTTCTATG	GAGGGAAATG	2100
	TGTGGTTTCC	TAGCTCTACA	GACATAACAG	CACAGCCCGA	TGTTGGATCA	GGCAGAGAGA	2160
	GCTTTCTCCA	GACTAATTAC	ACTGAGATAC	GTGTTGATGA	ATCTGAGAAG	ACAACCAAGT	2220
40	CCTTTTCTGC	AGGCCCAAGT	ATGTCACAGG	GTCCCTCAGT	TACAGATCTG	GAATGCCCAC	2280
	ATTATCTAC	CTTTGCTTAC	TTCCCAACTG	AGGTAACACC	TCATGCTTTT	ACCCCATCTT	2340
	CCAGACAACA	GGATTGGTCT	TCCACGGTCA	ACGTGGTATA	CTCGCAGACA	ACCCAACCGG	2400
	TATACAATGA	GGCCAGTAAT	AGTAGCCATG	AGTCTCGTAT	TGGTCTAGCT	GAGGGGTTGG	2460
	AATCCGAGAA	GAAGGCGAGT	ATACCCCTTG	TGATCGTGTC	AGCCCTGACT	TTTATCTGTC	2520
45	TAGTGGTTCT	TGTGGGTATT	CTCATCTACT	GGAGGAAATG	CTTCCAGACT	GCACACTTTT	2580
	ACTTAGAGTA	CAGTACATCC	CCTAGAGTTA	TATCCACACC	TCCAACACTT	ATCTTTCCAA	2640
	TTTCAGATGA	TGTGGAGCA	ATTCCAATAA	AGCACTTTCC	AAAGCATGTT	GCAGATTTAC	2700
	ATGCAAGTAG	TGGGTTTACT	GAAGAATTG	AGACACTGAA	AGAGTTTAC	CAGGAAGTGC	2760
	AGAGCTGTAC	TGTTGACTTA	GCTATTACAG	CAGACAGCTC	CAACCAACCA	GACAACAGC	2820
50	ACAAGAATCG	ATACATAAAT	ATCGTTGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	2880
	TGCTGAATA	GGATGGCAAA	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	2940
	ACAGACCAAA	AGCTTATATT	GCTGCCCAAG	GGCCACTGAA	ATCCACAGCT	GAAGATTTCT	3000
	GGAGAATGAT	ATGGGAACAT	AATGTGGAAG	TTATGTGTCAT	GATAACAAAC	CTCGTGAGAA	3060
	AAGGAAGGAG	AAAATGTGAT	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAC	3120
55	TTCTGGTCC	TCAGAAAGAT	GTGCAAGTGC	TTGCCCTATTA	TACTGTGAGG	AATTTTACTC	3180
	TAAAGAACAC	AGACTATAAA	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGA	CGTGTTGGTC	3240
	CACAGTATCA	CTACACGCG	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	3300
	TGACCTTTGT	GAGAAAGGCA	GCCTATGCCA	AGCGCCATGC	AGTGGGGCCT	GTGTGCTGCC	3360
	ACTGCAATGC	TGGAGTTGGA	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	3420
60	GAAATTCAACA	CGAAGGAAT	GTCAACATAT	TTGGCTTCTT	AAAAACATC	CGTTTACAAA	3480
	GAATTTATTT	GGTACAACAT	GAGGAGCAAT	ATGTCTTCAT	TCATGATACA	CTGGTTGAGG	3540
	CCATACCTTAG	TAAAGAACT	GAGGTGCTGG	ACAGTCATAT	TCATGCCAT	GTTAATGCAC	3600
	TCCCTATTCC	TGGACAGCA	GGCAAAACAA	AGCTAGAGAA	ACAAATCCAG	CTCCTGAGCC	3660
	AGTCAATAT	ACAGCAGAGT	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAGA	3720
65	ATCGAATTC	TTCTATCATC	CCTGTGGAAG	GATCAAGGAT	TGGCATTTCA	TCCCTGAGTG	3780
	GAGAAGGCC	AGACTACATC	AATGCCCTCCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	3840
	TCATCATTAC	CCAGCACCCT	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	3900
	ACCATAATGC	CCAATCGTGT	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	3960
	TTGTTTACTG	GCCAAATAAA	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACCTTTA	4020
70	TGGCTGAAGA	ACACAAATGT	CTATCTAATG	AGGAATAACT	TATAATTTCAG	GACTTTATCT	4080
	TAGAAGCTAC	ACAGATGAT	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	4140
	CAAATCCAGA	TAGCCCCATT	AGTAAACTTT	TTGAACCTAT	AAGTGTATTA	AAAGAAGAAG	4200
	CTGCCAATAG	GGATGGGCTT	ATGATTGTTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	4260
	CTTTCTGTGC	TCTGACAACC	CTTATGCACC	AACTAGAAAA	AGAAAATTC	GTGGATGTTT	4320
75	ACCAGGTAGC	CAAGATGATC	AATCTGATGA	GGCCAGGAGT	CTTGTCTGAC	ATTGAGCAGT	4380
	ATCAGTTTCT	CTACAAAGTG	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	4440
	CCACCTCTCT	GGACAGTAAT	GGTGACGAT	TGCCGTGATG	AAATATAGCT	GAGAGCTTAG	4500
	AGTCTTTAGT	TTAAACAGAA	AAGGGGTGGG	GGGACTCACA	TCTGAGCAIT	GTTTTCTCTT	4560
	TCCTAAATTT	AGGCAGGAAA	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTT	CCATCACCTG	4620
80	ACAGTAACCT	TCATGACATA	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	4680
	TTTTTGCAAG	ACTGTGAATT	TACTTATTAT	GTTTGAACCTA	AAATGATTGA	ATTTTACAGT	4740
	ATTTCTAAGA	ATGGAATTGT	GGTATTTTTT	TCTGTATTGA	TTTTTAACAGA	AAATTTCAAT	4800
	TTATAGAGGT	TAGGAATTCC	AAACTACAGA	AAATGTTTGT	TTTTAGTGTC	AAATTTTITAG	4860
	CTGTATTGT	AGCAATTATC	AGGTTTGCTA	GAATAATAAC	TTTTAATACA	GTAGCCCTGA	4920
	AATAAACAC	TCTTCCATAT	GATATTCAAC	ATTTTACAAC	TGCAATATTC	ACCTAAAGTA	4980
85	GAATAATCT	GTACTTATT	GTAAATACTG	CCCTAGTGTC	TCCATGGACC	AAATTTATAT	5040
	TTATAATTGT	AGATTTTAT	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATG	5100
	TTTAGTTTAA	TGACGTAGTT	CATTAGCTGG	TCTTACTCTA	CCAGTTTCT	GACATTGTAT	5160

TGTGTTACCT AAGTCATTAA CTTTGTTTCA GCATGTAATT TTAACCTTTG TGGAAAAATAG 5220
 AAATACCTTC ATTTTGAAG AAGTTTTTAT GAGAATAACA CCTTACCAAA CATTTGTTCAA 5280
 ATGGTTTTTA TCCAAGGAAT TGCAAAAAATA AATATAAATA TTGCCATTAA AAAAAAAA 5340
 AAAAAAAA AAAA AAAA

Seq ID NO: 186 Protein sequence:
 Protein Accession #: EOS sequence

1 11 21 31 41 51
 MVFKASKITP HWGKCNMSSD GSEHSLEGQK FPLEMQIYCF DADRFSSPFE AVKGKGLRA 60
 LSILFEVGTB ENLDFKAIID GVESVSRFGK QAALDPFILL NLLPNSTDKY YIYNGSLTSP 120
 PCTDTVDWIV FKDTVSISES QLAVFCEVLT MQQSGYVLM DLQNNFREQ QYKFSRQVFS 180
 SYTGKEEIHV AVCSSEPENV QADPENYTSI LVTWERPRV YDTMIEKPAV LYQQLDGEDQ 240
 TKHEFLTDGY QDLGAILNML LPNMSYVLQI VAICTNGLYG KYSDQLIVDM PTDNPELDLF 300
 PELIGTEBII KEEBEGKDIE EGAIVNPGRD SATNQIRKKE PQISTTHYN RIGTKYNEAK 360
 TNRSPTRGSE PSGKGDVPNT SLNSTSQPVT KLATEKDIDL TSQVTTELPP HTVEGTSASL 420
 NDGSKTVLRS PHMNLSTGAE SLNTVSITEY EESLSTSFK LDTGAEDSSG SSPATSAIPP 480
 ISENISQGYI FSSENPETIT YDVLIPESAR NASEDSTSSG SEESLKDPSM BGNVWFPSST 540
 DTTAQPDVGS GRESFLQNTY TEIRVDESEK TTKFSAGPV MSQGPSVTDL EMPHYSTFAY 600
 FPTEVTPHAF TPSSRQQLDV STVNVVYSQT TQPVYNEASN SSSESIRGLA EGLESEKQAV 660
 IPLVIVSALT FICLVVLVGI LIYWRKCPQT AHFYLEDSTS PRVISTPPTP IFPISDDVGA 720
 IPIKHFFPHV ADLHASSGFT EEPETLKEFY QEVQSCVDL GITADSSNHP DNKHKNRYIN 780
 IVAYDHSRVK SLQAEKDGK LTDYINANYV DGYNRPKAYI AAQGPLKSTA EDPWRMIWEH 840
 NVEIVIMTIN LVEKGRKKCD QYWPADGSEB YGNFLVTQKS VQVLAITYVR NFTLRNTKIK 900
 KGSQKGRPSG RVVTQYHYTQ WPDGMVPEYS LPVLTVPVKA AYAKRHAVGP VVHCASGVG 960
 RTGTIVVLDS FICLVVLVGI LIYWRKCPQT AHFYLEDSTS PRVISTPPTP IFPISDDVGA 1020
 EVLDSHIIHAY VNALIIPGPA GKTKLEKQFQ LLSQSNIQQS DYSAAKQCN REKNRTSSII 1080
 FVERSRVGIS SLQSGGTDTY NASYIMGYQ SNEFIITQHP LLHTIKDFWR MIWDHNAQLV 1140
 VMIPDQGNMA EDEFVYWPKN DEPINCESFK VTLMAEEHCK LSNEEKLIQ DFILEATQDD 1200
 YVLEVRHFQC PKWPNPDSPI SKTFELISVI KEAANRDGP MIVDEHGGV TAGTFCALIT 1260
 LMHQLEKENS VDVYQVAKMI NLMRPGVPAD IEQYQFLYKV ILSLVSTRQE ENPSTSLDSN 1320
 GAALPDGNIA ESLESLV

Seq ID NO: 187 DNA sequence
 Nucleic Acid Accession #: EOS sequence
 Coding sequence: 148-4632

1 11 21 31 41 51
 CACACATACG CACGACACGAT CTCACITCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAAAC ATTTCCCTCG CTCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
 CGGCGAGGGG CGCGAGACCG TCTGGAAATG CGAATCCTAA AACGTTTCCT CGCTTGCAIT 180
 CAGCTCCTCT GTGTTTGCCG CTTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTGTGTTAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAA TTGGGGAAAG 300
 AAATATCCAA CATGTAATAG CCCAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAAGGTT GGGATAAAAC ATCATTGGAA 420
 AACACATCA TCATATACAC TGGGAAACAA GTGGAAATTA ATCTCACTAA TGACTACCGT 480
 GTCAGCGGAG GAGTTTCAGA AATGGTGTIT AAAGCAAGCA AGATACTTT TCACTGGGGA 540
 AAATGCAATA TGTATCTGTA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600
 GAGATGCAAA TCTACTGCTT TGATGCGGAG CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660
 GGAAAGGAGG AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAAATTG 720
 GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTC GTTTTGGGAA GCAGCGCTGT 780
 TTAGATCCAT TCATCTGTTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTTAC 840
 AATGGCTCAT TGCATCTTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTGTGTG AAGTTCCTAC AATGCAACAA 960
 TCTGTGTTAG TCAATGCTGAT GGACTACTTA CAAAACAATT TTCGAGAGCA ACAGTACAA 1020
 TTCTCTAGAC AGGTGTTTTC CTATACACTT GGAAGGAGAG AGATTCTATG AGCAGTTTGT 1080
 AGTTCAAGAC CAGAAAAATG TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTGTGTACA 1140
 TGGGAAAGAC CTGAGTCTGT TTATGATACC ATGATTGAGA AGTTTGAGT TTTGTACCAG 1200
 CAGTGGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATTG TCAATAATTG GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380
 AATCCTGAAC TTGATCTTTT CCTGAATTA ATTGGAAGT AAGAAATAAT CAAGGAGGAG 1440
 GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500
 AACCATAACA GGAAGGAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACAGTCAC TAAATTAGCC 1680
 ACAGAAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAAGT TTCTTAGATC TCCACATATG 1800
 AACTTGTGCG GGACTGCAGA ATCCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860
 AGTTTATGTA CCAATTCTCA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920
 GCAACTCTG CTATCCCAAT CATCTCTGAG AACATATCCC AAGGTATAT ATTTCTCTCC 1980
 GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040
 GAAGATTCAA CTTCTACAGG TTCAGAGAA TCACTAAAGG ATCCTCTAT GGAGGGAAAT 2100
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGATC AGGCAGAGAG 2160
 AGCTTCTCTC AGACTAATTA CACTGAGATA CGTGTGTATG AATCTGAGAA GACAACCAAG 2220
 TCCTTTTCTG CAGGCCCACT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280
 CATTATTCTA CCTTGCCTTA CTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340
 TCCAGACACG AGGATTTTGT CTCCAGGTC AACGTGGTAT ACTGCAGAC AACCCACCG 2400
 GTATACATAG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGCTCTAGC TGAGGGGTTG 2460
 GAATCCGAGA AGAAGGCAGT TATACCCCTT GTGATCGTGT CAGCCCTGAC TTTTATCTGT 2520
 CTAGTGGTTC TTGTGGGTAT TCTCATCTAC TGGAGGAAAT GCTTCCAGAC TGCACACTTT 2580
 TACTTAGAGG ACAGTACATC CCCTAGAGTT ATATCCACAC CTCACACACC TATCTTTCCA 2640
 ATTTAGATG ATGTCCGAGC AATTCCAAATA AAGCACTTTC CAAAGCATGT TGCAGATTTA 2700
 CATGCAAGTA GTGGGTTTAC TGAAGAAAT GAGACACTGA AAGAGTTTTC CCAGGAAGTG 2760
 CAGAGCTGTA CTGTTGACTT AGGTATTACA GCAGACAGCT CCAACCAACC AGACAACAAG 2820

	CACAAGAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	2880
	CTTGTGTA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	2940
	AACAGACCAA	AAGCTTATAT	TGTCGCCCAA	GGCCCATCTGA	AATCCACAGC	TGAAGATTTC	3000
5	TGGGAATGA	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	3060
	AAAGGAAGGA	GAATATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	3120
	TTTCTGTGCA	CTCAGAAGAG	TGTGCAAGTG	CTTGCCCTATT	ATACTGTGAG	GAATTTTACT	3180
	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAAGGA	GACCCAGTGG	ACGTGTGGTC	3240
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	3300
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCATG	CAGTGGGGCC	TGTTGTGCTC	3360
10	CACGTGAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	3420
	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAAACACAT	CCGTTCCAAA	3480
	AGAAATTTAT	TGGTACAAAC	TGAGGAGCAA	TATGTCTTCA	TTCATGATAC	ACTGGTTGAG	3540
	GCCATACCTA	GTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTTATGCTTA	TGTTAATGCA	3600
	CTCCTCATT	CTGGACGAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GGGTCTCACT	3660
15	CTGTACACCA	GGCTGGAGTG	CAGAGGCACA	ATCTCGGCTC	ACTGCAACCT	TCCTCTCCCT	3720
	GGCTTAACAT	ATCCTCTCTAC	CTCAGCCTCC	CGAGTGGCTG	GGACTATACT	CCTGAGCCAG	3780
	TCAAATATAC	AGCAGAGTGA	CTATTCTGCA	GCCCTAAAGC	AATGCAACAG	GGAAAGAAGT	3840
	CGAACTCTCT	CTATCATCCC	TGTGGAAGA	TCAAGGGTTG	GCATTTCATC	CCTGAGTGGA	3900
	GAAGGCACAG	ACTACATCAA	TGCCTCCTAT	ATCATGGGCT	ATTACCAGAG	CAATGAATTC	3960
20	ATCATTACCC	AGCACCTCT	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	4020
	CATAATGCC	AAGTGGTGGT	TATGATTCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	4080
	GTTTACTTGG	CAATAAAGA	TGAGCCTATA	AATTGTGAGA	GCTTTAAGGT	CACCTCTATG	4140
	GCTGAAGAAC	CAAAATGTCT	ATCTAATGAG	GAATAACTTA	TAATTGAGGA	CTTTATCTTA	4200
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTTCAATGCT	TAAATGGCCA	4260
25	AATCCAGATA	GGCCCATTTG	TAAAACTTTT	GAACCTTATA	GTGTTATATA	AGAAGAAGCT	4320
	GCCAAATAGG	TGGGGCCTAT	GATTGTTTCT	GATGAGCATG	GAGGAGTGAC	GGCAGGAATC	4380
	TTCTGTGCTC	TGACCAACCT	TATGCACCAA	CTAGAAAAG	AAAATTCGCT	GGATGTTTAC	4440
	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	4500
	CAGTTTCTCT	ACAAAGTGAT	CCTCAGCCTT	GTGGGCACAA	GGCAGGAAGA	GAATCCATCC	4560
30	ACCTCTCTGG	ACAGTAATGG	TGCAGCATTG	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	4620
	TCCTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	4680
	CTAAATTTAG	CAGGAAAAAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCT	ATCACCTGAC	4740
	AGTAACTTTC	ATGACATAGG	ATCTGCGCGC	CAAATTTATA	TCATTAACAA	TGTGTGCTCT	4800
	TTTGCAAGAC	TGTAAATTTA	CTTATTATGT	TTGAACTAAA	ATGATTGAAT	TTTACAGTAT	4860
35	TTCTAAGAT	GGCAATTTGG	TATTTTCTTC	TGTATTGATT	TTAAACAGAA	ATTTCAATTT	4920
	ATAGAGGTTA	GGAAATCCAA	ACTACAGAAA	ATGTTTGTGT	TTAGTGTCAA	ATTTTATAGT	4980
	GTAATTTGAG	CAATTATCAG	GTTTGTCTAG	AATATAACTT	TTAATACAGT	AGCCTGTAAA	5040
	TAAACACTC	TTCCATATGA	TATTCACAT	TTTACAACCT	CAGTATTCAC	CTAAGTAGA	5100
	AATAATCTGT	TACTTATTTG	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	5160
40	ATAATTTGAT	TTTACTATCT	TTTACTACTG	AGTCAAGTTT	TCTAGTTCTG	TGTAATTTGT	5220
	TAGTTTAAAT	ACGTAGTTCA	TTAGCTGGTC	TTACTCTACC	AGTTTCTTGA	CATTGTATTG	5280
	TGTTTAAAT	CTTACTTAACT	TTGTTTCACT	ATGTAATTTT	AACCTTTTGT	GAAATAGAA	5340
	ATACTTCTAT	TTTGAAAGAA	GTTTTTATGA	GAATAACACC	TTACCAACAA	TTGTTCAAAT	5400
45	GGTTTTATC	CAAGGAATG	CAAAAATAAA	TATAATATT	GCCATTAAAA	AAAAAATAAA	5460
	AAAAAATAAA	AAAAAATAAA	A				

Seq ID NO: 188 Protein sequence:
Protein Accession #: EOS sequence

50	1	11	21	31	41	51	
	MRILKRPLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLKGL	KFGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSOGVSEMV	120
	FKASKITFMW	GKCNMSDGS	EHSLEGGKFP	LEMQIYCFDA	DRFSSFEEAV	KGKGLRLALS	180
55	ILFEVGTENN	LDFKAIIDGV	ESVSRFGKQA	ALDPPILLNL	LPNSTDKYIY	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISSQL	AVFCEVLTMQ	QSGYVMLMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGKEEIHAEV	CSSEPENVQA	DPENYTSLLV	TWERPRVVDY	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNLNLL	NMSYVLQIVA	ICTNGLYGYK	SDQLIVDMPT	DNPELDLFFE	420
	LIGTEETIKE	ESEGDIEEG	AIYNPGRDSA	TNQRKKEPEQ	ISTTHYNNRI	GTKYNEAKTN	480
60	RSPTRGSEFS	GKGDVPNTSL	NSTSQPVTKL	ATEKDILSLT	QTVTELPPTT	VEGTSASLND	540
	GSKTVLRSFH	MNLSTAESL	NTVSIETEEB	ESLLTSFKLD	TGAEDSGSSS	PATSAIPPIS	600
	ENISQGYIFS	SENPETITYD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NVWPPSSTDI	660
	TAQPDVGSGR	ESFLQNTYTE	IRVDESEKTT	KSFSAQPVMS	QGPSVIDLDM	PHYSTPAYFP	720
	TEVTPHAFPT	SSRQDLVST	VNVVYSQTQ	PVYNESANS	HESRIGLABG	LESEKKAIVP	780
65	LVIQSALTPI	CLVVLVGLLI	YWRKCFQTAH	FYLESTSPR	VISTPPTPIP	PISDDVGAIP	840
	IKHFFKHVAD	LEASSGPTTE	FETLKEFYQE	VQSCVVDLGI	TADSSNHPDN	KHKNNYINIV	900
	AYDHSRVKLA	QLAEXDGLIT	DIYNANYVDG	YNRPKAYIAA	QGPKLSTAE	FWRMIHEHNV	960
	EVIVMITNVL	EKGRKCDQY	WPADGSEYEG	NFLVTQKSVQ	VLAAYTVTRNF	TLRNTKIKKG	1020
	SQKGRPSGRV	VTQHYHTQWP	DMGVPEYSLP	VLTFRVKAAY	AKRHAVGPV	VHCSAGVGRT	1080
70	GTIIVLDSML	QQIQHEGTVN	IFGFLKHIRS	QRNYLVQTEE	QYVFIHDTLV	EAILLSKETEV	1140
	LDSSHAYVNV	ALLIPGPAGK	TKLEKQFQGL	TLSPRLCEGR	TISAHCNLPL	PGLTDPPTSA	1200
	SRVAGTILLS	QSNIQSDSYS	AALKQCNREK	NRTSSIIPEV	RSRVGSSLSL	GEGTDYINAS	1260
	YIMGYQSNB	PIITQHPLHL	TIKDFWRMIW	DHNAQLVVM	PDQNMMAEDE	FVYWPNKDEP	1320
	INCESPKVTL	MAEHEKCLSN	BEKLIQDFI	LEATQDDVVL	EVRRHQCPKW	PNPDSPISKT	1380
75	FELISVIKEB	AANRDGPMIV	HDEHGGVTAG	TPCALTTLMH	QLEKENSVDV	YQVARMINLM	1440
	RPGVFADIEQ	YQFLYKVLIS	LVGTRQENP	STSLDSNGAA	LPDGNIAESL	ESLV	

Seq ID NO: 189 DNA sequence
Nucleic Acid Accession #: NM_002820
Coding sequence: 304..831

85	1	11	21	31	41	51	
	COGGTTGSCA	AAGAAGCTGA	CTTCAGAGGG	GGAAACTTTC	TTCTTTTAGG	AGGCGGTTAG	60
	CCCTGTTCCA	CGAACCCAGG	AGAACTGCTG	GCCAGATTAA	TTAGACATTG	CTATGGGAGA	120
	CGTGTAACA	CACCTACTAT	CATTGATGCA	TATATAAAC	CATTTTATTT	TCGCTATTAT	180

5 TTCAGAGGAA GCGCCTCTGA TTTGTTCTT TTTTCCCTTT TTGCTCTTC TGGCTGTGTG 240
 GTTTGGAGAA AGCACAGTTG GAGTAGCCGG TTGCTAAATA AGTCCCGAGC GOGAGCGGAG 300
 ACGATGCAGC GGAGACTGGT TCAGCAGTGG AGCGTCGCGG TGTTCCTGCT GAGCTACGCG 360
 GTGCCCTCCT GCGGGCGCTC GGTGGAGGGT CTCAGCGCGC GCCTCAAAAG AGCTGTGTCT 420
 10 GAACATCAGC TCCTCCATGA CAAGGGGAAG TCCATCCAAG ATTTACGGCG ACGATTCTTC 480
 CTTCCACATC TGATCGCAGA AATCCACACA GCTGAAATCA GAGCTACCTC GGAGGTGTCC 540
 CCTAACTCCA AGCCCTCTCC CAACACAAAG AACCAACCCG TCGATTGG GTCTGATGAT 600
 GAGGGCAGAT ACCTAACTCA GGAACCTAAC AAGGTGGAGA CGTACAAAGA GCAGCCGCTC 660
 AAGACACCTG GGAAGAAAAA GAAAGGCAAG CCGGGAAAC GCAAGGAGCA GGAAAAGAAA 720
 15 AAAACGCGAA CTCGCTCTGC CTGGTTAGAC TCTGGAGTGA CTGGGAGTGG GCTAGAAGGG 780
 GACCACCTGT CTGACACCTC CACAACGTCG CTGGAGCTCG ATTCACGGTA ACAGGCTTCT 840
 CTGGCCCGTA GCCTCAGCGG GGTGCTCTCA GCTGGGTTTT GGAGCCTCCC TTCTGCCTTG 900
 GCTTGGACAA CCTAGAAAT TTCTCCCTTT ATGTATCTCT ATCGATTGTG TAGCAATTGA 960
 CAGAGAATAA CTCAGAATAT TGTCTGCCTT AAAGCAGTAC CCCCCTACCA CACACACCCC 1020
 20 TGTCTCTCCG CACCATAGAG AGGCGCTAGA GCGCATTCCT CTTTCTCCAC CGTACCCAA 1080
 CATCAATCCT TTACCATCTC ACCAAATAAT TTCATATTCA AGCTTCAGAA GCTAGTGACC 1140
 ATCTTCATAA TTTGCTGGAG AAGTGTATTT CTCCCTCTTA CTCTCACACC TGGGCAAACT 1200
 TTCTTCAGTG TTTTTCATT CTACGTTCT TTCACTTCAA GGGAGAATAT AGAAGCATTT 1260
 GATATTATCT ACAAACACTG CAGAACAGCA TCATGTCTAT AACGATTCTG AGCCATTAC 1320
 ACTTTTTATT TAATTAATTA TATTTAATTA AATCTCAAAT TTATTTTAAT GTAAAGAACT 1380
 TAAATTATGT TTTAAACACA TGCCTTAAAT TTGTTTAATT AAATTTAACT CTGGTTTCTA 1440
 CCAGCTCATA CAAAATAAAT GGTTCCTGAA AATGTTAAG TATTAACCTA CAAGGATATA 1500
 25 GGTTCCTCTC ATGTATCTTT TTGTTTCTTG GCAAGATGAA ATAATTTTTC TAGGTAATG 1560
 CCGTAGGAAA AATAAAACTT CACATTTAAA AAAAA

Seq ID NO: 190 Protein sequence:
Protein Accession #: NP_002811

30 1 11 21 31 41 51
 | | | | | |
 35 MQRRLVQWNS VAVFLLSYAV PSCGRSVEGL SRRLKRAVSE HQLLHDKGKS IQDLRRRFPL 60
 HHLIAEIHIA EIRATSEVSP NSKPSPNTRN HPVRFGSDDE GRYLTQETNK VETYKEQFLK 120
 TPGRKKGKPK GKKEQEKKK RRTRSWLDS GVTGSGLEGD HLSDTSTTSL ELDSR

Seq ID NO: 191 DNA sequence
Nucleic Acid Accession #: XM_059328
Coding sequence: 52..1023

40 1 11 21 31 41 51
 | | | | | |
 45 GGGCTGTCCG GCCCACTCCC CTGGGAGCGC GAGCGGTGGA CCGAGCGCGC CATGTCCCGC 60
 CCTCGCATGC GCCTGGTGGT CACCGCGGAC GACTTTGGTT ACTGCCCGCG ACGCGATGAG 120
 GGTATCGTGG AGGCCTTTCT GGC CGGGGCT GTGACCAAGG TGTCCCTGCT GGTCAACGGT 180
 GCGGCCACGG AGAGCGCGGC GGAGCTGGCC CGCAGGCACA GCATCCCGAC GGGCTCCAC 240
 50 GCCAACCTGT CCGAGGGCCG CCGCTGGGT CGCGCCCGCC GTGGCGCCTC ATCGCTGCTC 300
 GGCCCGGAAG GCTTCTTCTT TGGCAAGATG GGAATTCGGG AGGCGGTGGC GGCCGAGAGC 360
 GTGGATTTCG CTCAGGTGGG GGAGGAGCTC GAGGCCCAAC TAAGCTGCTT CCGGAGCTG 420
 CTGGGCAGGG CCCCCACGCA CGCGGACGGG CACCAGCAAG TGCAGTGTCT CCGAGCGGTG 480
 TGCCAGGTGT TCGCCGAGGC GCTGCAGGCC TATGGGGTGC GCTTTACGCG ACTGCCGCTG 540
 GAGCGCGGTG TGGGTGGCTG CACTTGGCTG GAGGCCCGCG CGCTGCTCTT CGCTCGCGCC 600
 GTGGAGCGGG ACGCCCGGGC CGCGGTGGGC CCGCTTCTCC GCCAGCGCCT GCGGTGGACA 660
 55 GACGCTTCTG TGGGCTGTAG CACTTGGCGC CGGCACATGT CCGCTCACCG CGTGTCCGGG 720
 GCCCTGGCGG GGTTCCTGGA AGGTACCTTA GCGGGCCACA CCGTGCAGCG CGAGCTGATG 780
 GCGCACCCCG GCTACCCAG TGTGCTCCC ACCGGCGGCT GGGTGAAGG CCGGACGCT 840
 TTCTCTTGCT CTTGGGAGCG GCTGCATGAG CTGCGCGTCC TCACCGCGCC CACGCTGCGG 900
 GCCCAGCTTG CCCAGGATGG CGTGACGCTT TGCGCCCTCG ACGACCTGGA CTCGAAGAGG 960
 60 CCAAGGGGAG AGGTCCCTCT TGAGCCCACT CTGGAACCTT TCCTGGAACC CTCCTACTCT 1020
 TGACCCCTTA CAGCAACACA AGCACTAATC CCGTTAGTAC CAAGAAAGGG GAGCCAGGAT 1080
 TTAGTCTCTG CCAAGCCAG AGCTGGGACC TGGAGCAGTA TCTGTGACT TCCTGGGTA 1140
 GGAACCTGCC ACCTCTGGGC TCAGGTCTCT ATGCTCTCAA ATGTCATCTA GAGTTTGAGC 1200
 AGCCTTCTTG GCTGAGGCA GGCCTAGCCT GTGGCAGCGG GCTAGGGCCC GCAGAGCAAT 1260
 65 TGGTGGCCCT CCAATGTGCA ATGCAACAC CTTCAACACT GGGGCGTGG GAGAGAGATG 1320
 CTATATTAAT AAAATAACGT GTGTCTTTC

Seq ID NO: 192 Protein sequence:
Protein Accession #: XP_059328

70 1 11 21 31 41 51
 | | | | | |
 75 MSRPRLV VADDFGYCPR RDEGIVEAFL AGAVTSVSLV VNGAATESAA ELARRHSIPT 60
 GLHANLSEGR PVGPARRGAS SLLGPEGFPL GKMGFREAVA AGDVLDPQVR EELEAQLSCF 120
 RELLEGRAPIH ADGHQHVHVL PGVCQVFAEA LQAYGVRETR LPLERGVGGC TWLEAPARAF 180
 ACAVERDARA AVGPFSRHGL RWTDAFVGLS TCGRHMSAHR VSGALARVLE GTLAGHTLTA 240
 ELMAHPGYPV VPPTGGCGEG PDAFSCSWER LHELEVLTA TLRQLAQDG VQLCALDLD 300
 SKRPGEEVPC EPTLEPFLP SL

Seq ID NO: 193 DNA sequence
Nucleic Acid Accession #: NM_005688.1
Coding sequence: 126..4439

80 1 11 21 31 41 51
 | | | | | |
 85 CCGGCGAGGT GGCTCATGCT CCGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCCGTGAGC 60
 AGGGGCGCAG GAATCTCTGAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCGCTCAG 120

	AGAAGATGAA	GGATATGAC	ATAGGAAAAG	AGTATATCAT	CCCCAGTCC	GGGTATAGAA	180
	GTGTGAGGGA	GAGAACACAG	ACTTCTGGGA	CGCACAGAGA	CGGTGAAGAT	TCCAAGTTCA	240
	GGAGAACTCG	ACCGTTGGAA	TGCCAAGATG	CCTTGGAAAC	AGCAGCCCGA	GCCGAGGGCC	300
5	TCTCTCTTGA	TGCTCTCATG	CATTCTCAGC	TCAGAACTCT	GGATGAGGAG	CATCCCAAGG	360
	GAAAGTACCA	TCATGGCTTG	AGTGCTCTGA	AGCCCATCCG	GACTACTTCC	AAACACCAAG	420
	ACCCAGTGA	CAATGCTGGG	CTTTTTCTCT	GTATGACTTT	TTCTGGGCTT	TCTTCTCTGG	480
	CCCGTGTGGG	CCACAAGAGG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCCAAG	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
10	AAGTTGGGCC	AGACGCTGCT	TCCCTGCGAA	GGGTTGTGTG	GATCTTCTGC	CGCACCAAGG	660
	TCATCTGTGC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAAGCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTAGT	GCTGGGCTTC	CTCCCTGACG	AAATCGTGGG	GTCTTGGTGG	CTTGCACTGA	840
	CTTGGGCAT	GAATTAACGA	ACCGGTGTCC	GCTTGGGGGG	GGCCATCCTA	ACCATGGCAT	900
15	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
	TTTGCTCCAA	CGATGGGCGA	AGAATGTTTG	AGGCAGCAGC	CGTTGGCAGC	CTGCTGCTCG	1020
	GAGGACCCGT	TGTTTGCCAT	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCCTCTCTGG	ATCAGCTGTT	TTTATCTCTT	TTTACCAGC	AATGATGTTT	GCATCACGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CGGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
20	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTTCTC	1260
	AGAGTGTTCA	AAAAATCCGC	GAGGAGGAGC	GTCCGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGATACAC	TGTGGGTGTC	GCTCCCATTT	TGGTGGTGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTCTATG	GACCTGGGCT	TTCCGATCTG	CAGCAGCACA	GGCTTTTACA	TGGTGGACAG	1440
	TCTTCAATTC	CATGACTTTT	GCTTTGAAAG	TAAACCCGTT	TTCACTAAAG	TCCCTCTCAG	1500
25	AAGCCTCAGT	GGCTGTTGAC	AGATTTAAGA	GTTTGTCTCT	AATGGAAGAG	GTTTCACTGA	1560
	TAAAGAACAA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAATAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCACTCCAGT	ATCCAGAATC	CGCCCAAGCT	GACCCCAAAA	ATGAAAAAAG	1680
	ACAAGAGGCG	TTCCAGGGGC	AAGAAAGAGA	AGGTGAGGCA	GCTGCAGGCG	ACTGAGCATC	1740
	AGGCGGTGCT	GGCAGAGCAG	AAAGGCCAOC	TCTCTCTGGA	CAGTGAACAG	CGGCCCAAGT	1800
30	CCGAAGAGGA	AGAAGGCAAG	CACATCCACC	TGGGCCACCT	GGCTTACAG	AGGACACTGC	1860
	ACAGCATCGA	CTCGAGATCA	CAAGAGGATA	AATCTGTTGG	AATCTGCGCG	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCAAT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGCAACCTTC	GCTTATGTGG	CCGACGAGCG	CTGATCTCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCTCTGTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
35	ACAGCTGCTG	CCTGAGGCTC	GACCTGGCCA	TTCTTCCCAG	CAGCGACCTG	ACGGAGATTG	2160
	GAGAGCGAGG	AGCCAACTCT	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCCT	GCCCGGGCCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCTCTG	ACGACCCCTC	CAGTGCCTTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGTCTTTC	ATGAAAGAGG	2400
40	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACGTATGAA	TTTAAATGGT	GACTATGCTA	2460
	CCATTTTAA	TAACTCTGTG	CTGGGAGAGA	CACGCCAGT	TGAGATCAAT	TCAAAAAAAG	2520
	AAACCCAGTG	TTCAAGAGAG	AAGTCAACA	ACAAGGGTCC	TAAACACAGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGGCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCACTGCC	CTGCTCAGTA	TATGCTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
45	TCCTGGTTAT	TATGGCCCTT	TTATGCTGTA	ATGTAGGCAG	CACCCGCTTC	AGCACCTGGT	2760
	GGTTGAGTTA	CTCGATCAAG	CAAGGAAGCG	GGAAACACAC	TGTGACTCGA	GGGAACGAGA	2820
	CTCGGTGAG	TGACAGCATG	AAGGACATC	CTCATATGCA	GTAATATGCC	AGCATCTAAG	2880
	CCCTCTCCAT	GGCCTCATG	CTGATCTCTG	AAGCCATTCG	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGCG	AGCTTCTCTC	CGGCTGCTAT	ACGAGCTTTT	CCGAAGGATC	CTTCAAGGCC	3000
50	CTATGAAGTT	TTTGAACAG	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAGAGCA	3060
	TGGATGAAGT	TGACGTGCGG	CTGCCGTTCC	AGGCCGAGAT	GTTTCTCCAG	AACGTTATCC	3120
	TGTTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCTCT	GTGGCAGTGG	3180
	GGCCCTCTGT	CATCTCTCTT	TCAGTCTCTG	ACATGTCTCT	CAGGGTCTCT	ATTCCGGAGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCACTCAC	CTTCTCTCTC	CCACATCACG	TCCAGCATAC	3300
55	AGGGCCTTGC	CACCATCCAC	GGCTACATAA	AAGGGCAGGA	GTCTCTGCAC	AGATACCAAG	3360
	AGCTGCTGGA	TGACAACCAA	GCTCTCTTTT	TTTGTGTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGGGGCT	GGACCTCATC	AGCATCGCCC	TCATCAACCA	CAGGGGGCTG	ATGATCGTTC	3480
	TTATGCAAGG	GCAGATTCCC	CCAGCCTATG	CGGCTCTGCG	CATCTCTTAT	GCTGTCCAGT	3540
	TAAAGGGGCT	GTTCAGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTACCTT	3600
60	CGGTGGAGAG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAAGCACT	GCCAGAATTA	3660
	AGAAACAAGG	TCCCTCCCTC	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGAGC	3720
	AGATGAGGTA	CGAGAAAGAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCTCTC	ACGATCAAA	3780
	CTAAAGAGAA	GATTGSCATT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTTCCG	CTGTGGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAATCA	3900
65	GTGATATTGG	CCTTGCCGAG	CTCCGAAGCA	AACCTCTTAT	CATTCTCTCA	GAGCCGGTGC	3960
	TGTTCACTGG	CACCTGTGAG	TCAAATTTGG	ACCCCTTCAA	CCAGTACACT	GAAGACGAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TTGAATCTGA	AGTGATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTATAGTAA	GCCACAGCTG	4200
70	CCATGGACAC	AGAGACAGAC	TTATGATTC	AAGAGACCAT	CCGAGAAGCA	TTTGCAAGCT	4260
	GTACCATGCT	GACCATTTGC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGGCCA	GGGACAGGTC	GTGGAGTTTG	ACACCCATC	GGTCTCTCTG	TCCAAGGACA	4380
	GTTCCCGATT	CTATGCCATG	TTTGCTGCTG	CAGAGAACAA	GGTCTGCTGC	AAGGGCTGAC	4440
	TCCTCCCTGT	TGACGAAGTC	TCTTTTCTTT	AGAGCAATTC	CATTCCCTGC	CTGGGGCGGG	4500
75	CCCTCATATG	GGTCTCTCTA	CGAAACCTTT	GCCTTTCTCG	ATTTTATCTT	TCCGACAGCA	4560
	GTTCCGAGTC	GGCTGTGCTG	TTTCACTTTT	AGGAGAGTCT	ATATTTTGAT	TATTGTATTT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTCTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATGT	TATCAGAGGC	CTATAATGAA	GCTTTATACC	TGTAGCTATA	4740
80	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
	TATTAAAAAT	AGCACTGTGC	TAAATAACAGT	GCATATCTCT	TTCTATCATT	TTTGTACAGT	4860
	TCTCTGTACT	AGAGATCTGG	TTTGTCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCTATCTT	4920
	CTCTAGCTGG	TGTTTTCAGG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCTCTGCTC	GGCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
	GAGAGCGGTC	GCCGCTGGGA	GACCATGTCAG	AGCGCCGTCG	GTCTCAGGG	CTCTGCTCTT	5100
	CTGTCTGTGG	GTCATCTACT	GTTTCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCTTC	5160
85	TCTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCTCT	TCTTCTTTTT	GCTGTGTTTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTCTT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340

5
10
15
20
25
30
35
40

```

GTTGGTTCCA AGCCCTGGAG CCAACTGCTG CTTTTTGAGG TGGCACTTTT TCATTGCTCT 5400
ATTGCCACAC CTCACAGATT CAGTGGCAGG GCTCAGGATT TCGTGGGTCT GTTTTCCTTT 5460
CTCACGCGAC TCGTGCACAC GTCTCTCTCT CTCTCTCCCC TCAAAGTCTG CAACTTTAAG 5520
CAGCTCTTTC TAATCAGTGT CTCACACTGG CGTAGAAGTT TTTGTACTGT AAAGAGACCT 5580
ACCTCAGGTT TCGTGGTTGCT GTGTGGTTTG GTGTGTTCCT GCAAAACCCC TTTGTGCTGT 5640
GGGGCTGGTA GCTCAGGTGG GGGTGGTCAC TGCCTGTCATC AGTTGAATGG TCAGCGTTGC 5700
ATGTCGTGAC CAACTAGACA TTCTGTGCGC TTAGCATGTT TGTGGAACAC CTTGTGGAAG 5760
CAAAAATCTG AAAATGTGAA TAAATTTATT TTGGATTGTT TAAAAAATAA AAAAAAATAA 5820
AAAAAATAA AAAAAATA

```

Seq ID NO: 194 Protein sequence:
Protein Accession #: NP_005679.1

15
20
25
30
35
40

```

1 11 21 31 41 51
MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRF TRPLECQDAL ETAARAEGLS 60
LDASMSHQLR ILDEEHPKKG YHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLLAR 120
VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQELNEV GPDAASLRV VWIFCRTL 180
LSIVCLMITQ LAGFSGPAPM VKHLLTYTQA TESNLQYSL LVLGLLLTET VRSWSLALTW 240
ALNYRTGVR LRGAILTMAFK KILKLKNIKE KSLGELINIC SNGDQRMPEA AAVGSLLAGG 300
PVVAILGMIV NVILIGPTGF LGSVAVFIFY PAMMFASRLT AYFRRCVAA TDERVQKMNE 360
VLTYIKFIKM YAVVKAFSOS VQKIREEERR ILEKAGYFQG ITVGVAPIVV VIASVVTFSV 420
HMTLGLDLTA AQAPTIVTVF NSMTFALKVT PFSVKSLSA SVAVDRPKSL FLMEEVHMIK 480
NKPASPHIKI EMQNATLAW SSSSIQNSP KLTFRMKDK RASRGKKEV RQLQRTHEQA 540
VLADQKGLHL LQSDERPSP EEEGKHILH HLRLQRTLHS IDLEIQEGL VIGCGSVGSG 600
KTSLSAILG QMTLLEGSIA ISGTAYVAQ QAWILNATLR DNILFGKEYD EERYNSVLNS 660
CCLRPDLAIL PSSDLTEIGE RQANLSGGQR QRISLARALY SDRSIYILD PLSALDAHVG 720
NHIFNSAIRK HLKSTKVLFFV THQLQYLVD DEVIFMKEG ITERGTHEEL MNLNGDYATI 780
FNHLLGETP PVEINSKKT SSGQKKSQDK GPKTGSVKKE KAVKPEEGQL VQLEKGGQS 840
VPWSVYGVYI QAAGGFLAPL VIMAFMLNV GSTAPSTWWL SYWIKQSGN TTVTRGNETS 900
VSDSMKDNPH MQYASIAIY SMAVMLILKA IRGVVFKGT LRASSRLHDE LFRRLRSPM 960
KFPDITPTGR ILNRFSKMD EVDVRLPFA EMFIQNVILV FFCVGMIAV PFWFLVAVGP 1020
LVILFSLVHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNK QEFLHRYQEL 1080
LDNQAPFPL FTCAWRWLA RDLISIALI TTTGLMIVLM HGQIPPAYAG LAISYAVQLT 1140
GLEQFTVRLA SETEARFTSV ERINHYIKL SLEAPARIK KAPSPDWQGE GEVTFENAFM 1200
RYRENLPVLV KKVSPFIKX EKIGIVRGT GSGKSLGML FRLVELSGG IKIDGVRISD 1260
IGLADLRSLK SIIPQEPVLF SGTVRSNLD PNOYTEDQIV DALERTHKE CIAQLPLKLE 1320
SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIQE TIREAFADCT 1380
MLTIAHRLIT VLGSDRIMVL AQGGVVEFDT PSVLLSNDSS RPYAMFAAA NKVAVKG

```

Seq ID NO: 195 DNA sequence
Nucleic Acid Accession #: NM_006470
Coding sequence: 228..1922

45
50
55
60
65
70
75
80
85

```

1 11 21 31 41 51
GCTGTCTGCA GCTCTAGTAC TCTAGCTGCC TTGTGCGCAT CGCATCTGGC TGCCATCCAG 60
CGCCAGCACA CAGTAATAG TGGCCGAGCT TCCTCTGGGA GGGAGGAAAC AGTTAAATC 120
TTGCAGCAGC TGCAATCATC TAGGCGTGGT TCTCTGTCT GACTTGGGCT GCACAGATCC 180
TGGGCCAAGG GACAGAGAA AGACAGCCTA GGAGCAGAGC CTCCAGATG GCTGAGTTGG 240
ATCTAATGCG TCCAGGCCA CTGCCCAGGG CCACCTGCTCA GCCCCAGGCC CCTCTCAGCC 300
CAGACTCTGG GTACCCAGC CCAGATTCTG GGTGAGCCAG CCCAGTGGAA GAAGAGGAGC 360
TGGGCTCCTC GGAGAAGCTT GGCAGGGAGA CGGAGGAACA GGACAGCGAC TCTGCAGAGC 420
AGGGGGATCC TGCTGCTGAG GGGAAAGAGG TCCTGTGTGA CTCTGCTT GATGACACCA 480
GAAGAGTGAA GGCAGTGAAG TCCTGTCTAA CCTGCATGCT GAATTAAGT GAAGAGCACT 540
TGCAAGCCGA TCAGGTGAAC ATCAAACTGC AAAGCCACCT GCTGACCGAG CCAAGTGAAG 600
ACCAACATG GCGTACTGCT CCTGCCACAC ACAGCCACT GTCTGCTTTC TGCTGCCCTG 660
ATCAGCAGTG CATCTGCCAG GACTGTTGCC AGGAGCACAG TGGCCACACC ATAGTCTCCC 720
TGATGTCAGC CCGCAGGAC AAGGAGGCTG AACTCCAGTG CACCCAGTTA GACTTGGAGC 780
GGAAACTCAA GTTGAATGAA AATGCCATCT CCAGGCTCCA GGCTAACCA AAGTCTGTT 840
TGGTGTGCGT GTCAGAGGTC AAAGCGGTGG CTGAATGCA GTTTGGGGAA CTCCTTGCTG 900
CTGTGAGGAA GGCCAGGCC AATGTGATG TCTTCTAGA GGAGAAGGAG CAAGCTGGCC 960
TGAGCCAGGC CAACGGTATC AAGGCCACCT TGGAGTACAG GAGTGCCGAG ATGAGAGAAG 1020
GCAAGCAGGA GCTGGAGAG ATGCGCGCCA TCAGCAACAC TGTCCAGTTC TTGAGGAGT 1080
ACTGCAAGTT TAAGAACACT GAAGACATCA CCTTCCCTAG TGTTCAGTA GGCCTGAAG 1140
ATAAACTCTC GGGCATCCGC AAAGTTATCA CGGAATCCAC TGTAACATTA ATCCAGTTGC 1200
TGGAAACTA TAAGAAAAG CTCCAGGAGT TTTCCAAGGA AGAGGAGTAT GACATCAGAA 1260
CTCAAGTGTG TGGCTGTGTT CAGCGCAAAT ATTGGACTTC CAAACCTGAG CCCAGCACCA 1320
GGGAACAGTT CCTCCAATAT GCGTATGACA TCACTGTTGA CCGGACACA GCACACAAGT 1380
ATCTCOGGCT GCAGGAGGAG AACCGCAGG TCACCAACAC CACGCCCTGG GAGCATCCCT 1440
ACCGGACCT CCGCAGCAGG TTCCTGCACT GCGGCGAGGT GCTGTCCAG CAGAGTCTGT 1500
ACCTGCACAG GTACTATTTT GAGGTGGAGA TCTTGGGGC AGGCACCTAT GTTGGCCTGA 1560
CCTGCAAGG CATGACCGG AAAGGGGAG AGCGCAACAG TTGCATTTCC GGAAACCACT 1620
TCTCTGGAG CTCTCAATGG AACGGGAAGG AGTTCAAGGC CTGTACAGT GACATGGAGA 1680
CCCCACTCAA AGCTGGCCCT TTCGGAGGCT TCGGGGCTA TATCGACTTC CCGGAGGGA 1740
TCTTTTCTCT CTATGGCGTA GAGTATGATA CCATGACTCT GGTTCACAG TTTGCTGCA 1800
AATTTTCAGA ACCAGTCTAT GCTGCTTCT GGTCTTCAA GAAGGAAAAC GCCATCCGGA 1860
TTGTAGATCT GGGAGAGGAA CCGGAGAAGC CAGCACCGTC CTGGGGGTG ACTGCTCCCT 1920
AGACTCCAGG AGCCATATCC CAGACCTTTG CCAGCTACAG TGATGGGATT TGCAATTTAG 1980
GGTGATTGT GGGCAGAAAT AACTGCTGAT GGTAGCTGGC TTTTGAAATC CTATGGGGTC 2040
TCTGAATGAA AACATCTTCC AGCTGCTCTC TTTTGTCCA TATGGTGTCT TTCTCTATG 2100
GTTTGCAGTA ATTCTTTT TTTTTTTGA GACGGAGTCT CGCACTGTG CCCAGGCTGG 2160
AGAGCAGTGG CGCGATCTTG GCTCACTGCA AGCTCCGCT CCGAGTTCA AGCAATTCTC 2220
CTGCTCAGC CTCCGAGTA GCTGGGATTA CAGGTGCTG CCACACAC CAGCTAATGT 2280
TTGTATTTT TAGTAGAGAT GGGGTTTCAC CATGTTGGCC AGGCAGATCT CAAACTCTCT 2340

```

ACCTCGTGAT GCACCCACCT CGGCCTCCCA AAGTGCTGGG ATTACATGCG TGAGCCACTG 2400
 CGCCCTGCCT GTTTGTAGTA ATTTTITAGG ACCAAATCTC CCTCATCTTC TAGTGCCATT 2460
 CTCTCTCTCG TTCAGGTAAA TGTCACTG TGCCCAAGAT GGATGACCAG GAACCTTAAA 2520
 GAGTGGCTGA AAAGATTGCA GAGTTATCAT AATAAATTGC TAACTTGGCT

5

Seq ID NO: 196 Protein sequence:
 Protein Accession #: NP_006461

10

1 11 21 31 41 51
 MAELDLMAPG PLPRATAQPP APLSPDSGSP SPDSGSASPV EEDVGSSEK LGRETEEQDS 60
 DSABQGDPAQ EGKEVLCDFC LDDTRRVKAV KSCLTCMVNY CEEHLQPHQV NIKLQSHLLT 120
 EPVKDHNWRY CPAHHSPLSA FCCPDQQCIC QDCCQEHSGH TIVSLDAARR DKEAELQCTQ 180
 15 LDLERKLLKN ENAISRLQAN QKSVLVSVSE VKAVAEMQFG ELLAAVRKAQ ANVMLFLEEK 240
 EQAALSQANG IKAHLEYRSA EMEKSKQELE RMAAISNTVQ FLEBYCKPKN TEDITFPSPVY 300
 VGLKDLKSI RKVITESTVH LIQLLENYKK KLQEFSEKEE YDIRTQVSAV VQRKYWTSKP 360
 EPSTRBQFLQ YAYDITDPD TAHKYLRLQE ENRKVINTTP WEHPYDLPVS RFLHWRQVLS 420
 20 QOSLYLHRY FEVRIFGAGT YVGLTCKGID RKGEERNSCI SGNFNSWSLQ WNGKEFTAMY 480
 SDMETPLKAG PFRRLGVYID FPGGILSFYG VEYDTMTLVH KFAKFPSEPV YAAFWSLKKE 540
 NAIRIVDLGE EPEKPAPSLG VTAP

20

Seq ID NO: 197 DNA sequence
 Nucleic Acid Accession #: NM_004316
 Coding sequence: 433-1149

25

1 11 21 31 41 51
 CCCGAGACCC GCGCGAAGAG AGCGCAGCCT TAGTAGGAGA GGAACCGCAG ACGCGGCAGA 60
 30 GCGGGTTCTAG CACTGACTTT TGCTGCTGCT TCTGCTTTT TTTTCTTAG AAACAAGAAG 120
 GCGCCAGCGG CAGCCTCACA CGCGAGCGCC ACGCGAGGCT CCGAAGCCA ACCCGCGAAG 180
 GGAGGAGCGG AGGAGGAGG AGGCGGCGTG CAGGGAGGAG AAAAAGCATT TTCACCTTTT 240
 TTGCTCCAC TCTAAGAAGT CTCCCGGGGA TTTGTATAT ATTTTAAAC TTCGTCAGG 300
 35 GCTCCCGCTT CATATTCTCT TTTCTTTCCC TCTCTGTTC TGACCCCAAG TTCTCTCTGT 360
 GTCCCGCTCG CGGCGCCGCG ACCTCGCGTC CGGATCGCT CTGATTCGCG GACTCCTTGG 420
 CCGCGCTGCG GCATGGAAG CTCTGCCAAG ATGGAGAGCG GCGGCGCGG CCGAGCAGCCC 480
 CAGCGCAGCT CCGCAGCAGC CTCTCTGCG CCGCAGCCT GTTCTTTGC CAGCGCGCA 540
 GCGCGCGCGG CCGCAGCGCG CGCAGCGGCA GCGCAGAGCG CGCAGCAGCA GCAGCAGCAG 600
 40 CAGCAGCAGC AGCAGCAGCA GCAGGCGCGG CAGCTGAGAC CGGCGGCGGA CGGCCAGGCC 660
 TCAGGGGCGG GTCACAAGT AGCGCCCAAG CAAGTCAAGC GACAGCGCTC GTCTTCGCCC 720
 GAACCTGATCG GCTGCAAAAG CCGGCTCAAC TTCAGCGGCT TTGGCTACAG CCGTCCGCGAG 780
 CAGCAGCGCG CGCGCGTGGC GCGCGCAAC GAGCGGAGC GCAACCGCGT CAGATTGGTC 840
 AACCTGGGCT TTGCCACCTT TCGGAGCAC GTCCCCAAG GCGCGGCCAA CAAGAAGATG 900
 45 AGTAAGGTGG AGACACTGCG CTGCGCGGTC GAGTACATCC GCGCGCTGCA CGAGCTGCTG 960
 GACGAGCATG ACGCGGTGAG CGCGCGCTTC CAGGCGGCG TCCTGTGCGC CACCATCTCC 1020
 CCCAATCTACT CCAAGCACTT GAACTCCATG CGCGGCTCGC CGGTCTCATC CTACTCGTGG 1080
 GACGAGGCGT CTTAAGACCC GCTCAGCCCC GAGGAGCAGG AGCTTCTGCA CTTCAACCAAC 1140
 TGGTCTGAG GGGCTCGGCC TGCTCAGGCC CTGGTTCGAA TGGACTTTGG AAGCAGGGTG 1200
 50 ATCGCACAC CTGCATCTTT AGTGCTTTCT TGTCAGTGGC GTTGGGAGGG GGAGAAAAGG 1260
 AAAAGAAAAG AAAGQPSGGG HKSAPKQVQR QRSSPELMR CKRRLNFSGF GYSLPQQQPA 1320
 CCAACCCCAT CGCCCACTAA CGGAGGCATG CCTGAGAGAC ATGGCTTTCA GAAACCGGGA 1380
 AGCGCTCAGA ACAGTATCTT TGCACCTCAA TCAITCACGG AGATATGAAG AGCAACTGGG 1440
 ACCTGAGTCA ATGCGCAAAA TGCAGCTTGT GTGCAAAAGC AGTGGGCTCC TGGCAGAAGG 1500
 55 GAGCAGCACA CGCGTTATAG TAACTCCCAT CACCTCTAAC ACGCAGAGCT GAAAGTTCTT 1560
 GCTCGGCTCC CTTCACCTCC CGGCCCTTTC TTAGAGTGCA GTTCTTAGCC CTCTAGAAAC 1620
 GAGTTGGTGT CTTTC

55

Seq ID NO: 198 Protein sequence:
 Protein Accession #: NP_004307

60

1 11 21 31 41 51
 MESSAKMESG GAGQPPQPP QPPFLPPAAC FFATAAAAAA AAAAAAQSA QQQQQQQQQQ 60
 65 QQQQAPQLRP AADQPPSGGG HKSAPKQVQR QRSSPELMR CKRRLNFSGF GYSLPQQQPA 120
 AVARRNERER NRKLVNLGF ATLREHVPNG AANKKMSKVE TLRSAVEYIR ALQQLLDEHD 180
 AVSAAPQAGV LSPTISPNYS NDLSNMGASP VSSYSDEGS YDPLSPDEQE LLDFTNWF

65

Seq ID NO: 199 DNA sequence
 Nucleic Acid Accession #: NM_007015
 Coding sequence: 1-1005

70

1 11 21 31 41 51
 75 ATGACAGAGA ACTCCGACAA AGTTCCTCATT GGCCTGGTGG GACCTGATGA CGTGAATTTC 60
 TGCAGCCCCC CGCGGTACGC TACGCTGACG GTGAAGCCCT CCAGCCCGCG GCGGCTGCTC 120
 AAGGTGGGAG CCGTGTCTCT CATTTCCGGA GCTGTGCTGC TGCTCTTTGG GGCCATCGGG 180
 GCCTTCTACT TCTGGAGGG GAGCGACAGT CACATTTACA ATGTCCATTA CACCATGAGT 240
 80 ATCAATGGGA AACTACAAGA TGGGTCAATG GAAATAGACG CTGGGAACAA CTTGGAGACC 300
 TTTAAATGG GAAGTGGAGC TGAAGAAGCA ATTGCAGTTA ATGATTTCOA GAATGGCATC 360
 ACAGGAATTC GTTTTGTCTG AGGAGAGAAG TGCTACATTA AAGCGCAAGT GAAGGCTCGT 420
 ATTCTGTAGG TGGGCGCGGT GACCAACAG AGCATCTCCT CCAAACTGGA AGGCAAGATC 480
 ATGCCAGTCA AATATGAAGA AAATTCTCTT ATCTGGGTGG CTGTAGATCA GCCTGTGAAG 540
 85 GACAACAGCT TCTTGTGTTT TAAGGTGTTA GAACTCTGCG GTGACCTTCC TATTTTCTGG 600
 CTTAAACCAA CCTATCCAAA AGAAATCCAG AGGGAAAGAA GAGAAGTGGT AAGAAAAATT 660
 GTTCCAACTA CCACAAAAAG ACCACACAGT GGACCAACGA GCAACCCAGG CGCTGGAAGA 720
 CTGAATAATG AAACAGAGAC CAGTGTTCAG GAGGACTCAC AAGCCTTCAA TCCTGATAAT 780

85

CCTTATCATC AGCAGGAAGG GGAAAGCATG ACATTGACG CTAGACTGGA TCACGAAGGA 840
 ATCTGTTGTA TAGAATGTAG GCGGAGCTAC ACCCACTGCC AGAAGATCTG TGAACCCCTG 900
 GGGGGCTATT ACCCATGGCC TTATAATTAT CAAGGCTGCC GTTCGGCCTG CAGAGTCATC 960
 ATGCCATGTA GCTGGTGGGT GGCCCGTATC TTGGGCATGG TGTGAAATCA CTTCATATAT 1020
 CACGTGCTGT AAAATAAGAA CTAGCTGAAG AGACAACCAA AGAAGCATTA AGGCAGGTG 1080
 ATGCTGATGG GACCATAAAA TATTTTACA CGCAGCCTGA GCGGTTATTC TTGACACTCT 1140
 TAACAGAAAT TTTTAAATCG TTTTCCAGAA CTTAGTATA TGCAAAATGA CTGAAAGGGT 1200
 AGTTCAGATC TAAATGCGCA TAACCCCGTT ATTTGTTATT TTTTATTGCG ATTGATTGCG 1260
 CATAAGTCTT CCCTTGCTTG CATCTTCCAA AGCTATTTCG AAATAAACAC GAAAATTTAC 1320
 AGTTTGCC

Seq ID NO: 200 Protein sequence:
 Protein Accession #: NP_008946

1 11 21 31 41 51
 MTENSDKVPF ALVGPDDVEF CSPPAYATLT VKPSSPARLL KVGAVVLISG AVLLLPFAIG 60
 AFYFWKSGDS HIYNVHYTMS INGKLQDQSM BIDAGNNLET PKMGSGAEEA IAVNDPQNGI 120
 TGIRFAGGEK CYIKAQVKAR IPEVGAVTKQ SISKLEGGKI MPVKYEENSL IHWAVDQPVK 180
 DNSFLSSKVL ELCGLPIFW LKPTYPKEIQ RERREVVVKI VPTTTKRPHS GPRSNPAGGR 240
 LNNETRPVSQ EDSQAFNPDN PYHQQEGESM TFDPRLDHEG ICCIECRRSY THQKICEPL 300
 GGYYPWPYNY QGCRSACRVI MPCSWWVARI LGMV

Seq ID NO: 201 DNA sequence
 Nucleic Acid Accession #: NM_000728.2
 Coding sequence: 112..495

1 11 21 31 41 51
 GTAATAAGAG CCGGGTCTCC GCGGGGAAGG CGCCACAGC AGGTGTGGTG TTCATCCCGG 60
 GTCGACCGGC CGCTCGCGCT GCGCTGAAAC TCTAGTCGCC AGAGAGGCGG CATGGGTTTC 120
 CGGAAGTTCT CCCCCCTCCT GGCTCTCAGT ATCTTGGTCC TGTACCAAGC GGGCAGCCTC 180
 CAGGCGGCGC CATTCAAGTC TGCCCTGGAG AGCAGCCGAG ACCCGGCCAC ACTCAGTAAA 240
 GAGGACGCGC GCCTCTCTGT GGCTGCACTG GTGCAGGACT ATGTGCAGAT GAAGGCCAGT 300
 GAGCTGAAGC AGGAGCAGGA GACACAGGGC TCCAGCTCCG CTGCCAGAA GAGAGCCTGC 360
 AACACTGCCA CCTGTGTGAC TCATCGGCTG GCAGGCTTGC TGAGCAGATC AGGGGGCATG 420
 GTGAAGAGCA ACTTCGTGCC CACCAATGTG GGTTCCAAAG CCTTTGGCAG GCGCCGCAGG 480
 GACCTTCAAG CCTGAGCAGA TGAATGACTC CAGGAAGAAG GTGTGTCTTA AATCCAATGA 540
 CATATCCTTA TAAGAGATTG ACTCAGAAGA CACATGTGGA GAAGGTGACA TGACAGAGGC 600
 AAGGAGGCAC AAGCCAAGGA AGTCTGTGTC TACCAGAAGC CAGAATCACA GAACAGTCTC 660
 TGGAGAAGA GCAGCCCTGC TGACACCTAG AGTTTGGACT TCCAGCTTCC AGAAGTGTGA 720
 GAGAATAATT TCTGTTGTTT TAAGCCACAA AGTTTGTGGT AATTTGTTAT GACAGCCCTA 780
 GGAACATAAT ACAATACATT TTCAATTTAT TTGGGTAAAT GCCTTGGAGT GGGATTGCTG 840
 GGTTAATTTG AAGTGTGTTA TTTAACTCTG TAAGAAACTG CCAAACTATT TTCTGAAGTG 900
 ACTGTACCAC TTGCGCTTCT TGCCAGCCAC ATATGAGAGC TCTAGTATTT CCACAAATAG 960
 GTATGTAGCA GTAICTCAIT GCTGTTTTAA TTTGTATTTC CCAATGACT AATGACGTTG 1020
 AGCATCTATT TTACCATATG TTTATCACTT TTATTGAAGG GTCTGTTTAA ATCTTCTGCT 1080
 AAATTTTGTG TGGCTTGCTT GCTTTATTAG TGTGTAGTTT TTAGAGCTCT TTATATGTTG 1140
 TGGATGCAAG ATTGTTTCTA GATATATAGT TTGGAAACTT CCTTCCCTCG AATCTGCGGA 1200
 TTGCTTTTTC ATTTTCTTAG CAGTGTCTCT CACAGAGAAA AAGTTGTAAT TTGAATAAGA 1260
 TCCAATTCAT CTTTTTTTTT CTTTTATGTA TTGTGCTTTT AGTTCACTGC TAAGAACTCT 1320
 TTGCTAACT AAGGTCCCAA GGTCAACAATA ACCTTATTCT ATACTTTCTT GTAAAAGTTT 1380
 TATAGTTTAA TATTTTATAT GTAGATTAGT GATCTATTTT GAGTTAATTT TTGTATAAGG 1440
 TGAGAGGTGT AGGTGGAAT TCATACCTGT GAATATAGAT ACCCAATTGT TTCAGTGCCA 1500
 TTTGTTAAAA AGACTGTTAT TTCACCATTT AATTGCCCTT GCACCTTGT CAAAAGCAA 1560
 CTGATCATAT TTGTGTGGGT ATATTCTGCG GTTCTCAATT CRTCTCATT GATTGATTG 1620
 ACCATTCTTT TGCCAATGTC ATACTGCTT GATTAGTGA GTGTAAAGT GAATCTCAAA 1680
 ACCAGATAAT GTGGCTCTAC CAACATGTTT CATCTTGTG CAAAAGATT TTAGCTACAT 1740
 CTAAATATT TTCTACATCT TTTATACATT TTAGAATCAG TGTGTTACTA TCTACAAAT 1800
 TTCTGATGAG ATTTTAAATG GGATTGTGTT AAATCAGTGG GTTAATTTTG GGAGAAATTAG 1860
 CATATTAATA ATATTAAGTC GTTCAATTCA TGAACACAA ACATGTTTTT ACTTATTAG 1920
 GTTTCTCTG TTTTTTTTTT TTTAACAGTG TTCTCAGTTT TCAACAGAAA TATTCTACAC 1980
 ATATCTGTT AGATTTTAA CTATTTTATT TTTTGGTCT AATGTAAATG GTACTTAAAC 2040
 ATTTTGTGTT TTAATGTGTC ATTGCTAGTA GATAGAATA CAATATTTAA AATATTAGGA 2100
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 202 Protein sequence:
 Protein Accession #: NP_000719.1

1 11 21 31 41 51
 MGFRKFSFPL ALSILVLYQA GSLQAAPFRS ALESSPDPAT LSKEDARLLL AALVDYVQM 60
 KASELKQEQE TQSSSSAAQK RACNTATCVT HRLAGLLSRG GGMVKSFPV TNVGSKAFGR 120
 RRRDLQA

Seq ID NO: 203 DNA sequence
 Nucleic Acid Accession #: NM_001741
 Coding sequence: 71..496

1 11 21 31 41 51
 CTCTGGCTGG ACGCGCGCGC CGCCGCTGCC ACCGCTCTG ATCCAAGCCA CCTCCCGCCA 60
 GAGAGGTGTC ATGGGCTTCC AAAAGTTCTC CCCCTTCTG GCTCTCAGCA TCTTGTCTCT 120
 GTTGCAAGCA GGCAGCTTCC ATGCAGCACC ATTCAGGTCT GCCCTGGAGA GCAGCCGAGC 180
 AGACCCGGCC ACGCTCAGTG AGGACGAAGC GCGCTTCTG CTGGCTGCAC TGGTGCAGGA 240
 CTATGTGCAG ATGAAGGCCA GTGAGCTGGA GCAGGAGCAA GAGAGAGAGG GCTCCAGCCT 300

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

```

GGACAGCCCC AGATCTAAGC GGTGCGGTAA TCTGAGTACT TGCATGCTGG GCACATACAC 360
GCAGGACTTC AACCAAGTTTC ACACGTTCCC CCAAACCTGCA ATTGGGGTTG GAGCACCTGG 420
AAAGAAAAGG GATATGTCCA GCGACTTGGA GAGAGACCAT CGCCCTCATG TTAGCATGCC 480
CCAGAAATGCC AACTAAACTC CTCCCTTTCC TTCTTAATTT CCCTTCTTGC ATCCTTCTTA 540
TAACTTGATG CATGTGGTIT GGTTCCTCTC TGGTGGCTCT TTGGGCTGGT ATTGGTGGCT 600
TTCTTGTGGG CAGAGGATGT CTCAAACCTC AGATGGGAGG AAAGAGAGCA GGACTCACAG 660
GTTGGAAGAG AATCAGCTGG GAAAATACCA GAAATGAGG GCCGCTTTGA GTCCCCCAGA 720
GATGTCATCA GAGCTCCTCT GTCCTGCTTC TGAATGTGCT GATCAATTGA GGAATAAAAT 780
TATTTTTCCC C
  
```

Seq ID NO: 204 Protein sequence:
Protein Accession #: NP_001732

1 11 21 31 41 51
 MGPKQKSPFL ALSILVLLQA GSLHAAPFRS ALESSPADPA TLSEDEARLL LAALVQDYVQ 60
 MKASELEQEQ ERREGSLDSP RSKRCGNLST CMLGTYTQDF NKPHTFPPQTA IGVGAPGKKR 120
 DMSSDLERDE RPHVSM PQNA N

Seq ID NO: 205 DNA sequence
Nucleic Acid Accession #: NM_005361
Coding sequence: 1-945

1 11 21 31 41 51
 ATGCCTCTTG AGCAGAGGAG TCAGCACTGC AAGCCTGAAG AAGGCCTTGA GGCCCGAGGA 60
 GAGGCCCTGG GCTGTGTGGG TGCGCAGGCT CCTGCTACTG AGGAGCAGCA GACCGCTTCT 120
 TCCTCTTCTA CTCTAGTGGG AGTTACCCCTG GGGGAGGTGC CTGCTGCCGA CTCACCGAGT 180
 CCTCCCCACA GTCTCTCAGG AGCCTCCAGC TTCTCGACTA CCATCAACTA CACTCTTTGG 240
 AGACAATCCG ATGAGGGCTC CAGCAACCAA GAAGAGGAGG GGCCAAGAAT GTTTCCCGAC 300
 CTGGAGTCCG AGTTCCAAGC AGCAATCAGT AGGAAGATGG TTGAGTTGGT TCATTTTCTG 360
 CTCCTCAAGT ATCAGCCAGG GGAGCCCGTC ACAAAGGCAG AAATGCTGGA GAGTGTCTCT 420
 AGAAATGGCC AGGACTTCTT TCCCGTGATC TTCAGCAAGG CCTCCGAGTA CTTGCAGCTG 480
 GTCTTTGGCA TCGAGGTGGT GGAAGTGGTC CCCATCAGCC ACTTGTACAT CCTTGTCAAC 540
 TGCTGTGGCC TCTCTACAGA TGGCCTGCTG GGCAGACAATC AGGTCTATGCC CAAGACAGGC 600
 CTCCTGATAA TCGTCTCTGG CATAATGCA ATAGAGGGCG ACTGTGCCCC TGAGGAGAAA 660
 ATCTGGGAGG AGCTGAGTAT GTTGAGGTG TTTGAGGGGA GGGAGGACAG TGTCTTCCGA 720
 CATCCAGGAA AGCTGCTCAT GCAAGATCTG GTGCAGGAAA ACTACCTGGA GTACCGGCAG 780
 GTGCCCGGCA GTGATCTCTG ATGCTACGAG TTCTGTGGGG GTCCAAGGGC CCTCATTGAA 840
 ACCAGCTATG TGAAGTCTCT GCACCATACA CTAAGATCG GTGGAGAAC TCACATTTC 900
 TACCCACCCC TGCATGAACG GCCTTTGAGA GAGGAGAGAG AGTGA

Seq ID NO: 206 Protein sequence:
Protein Accession #: NP_005352

1 11 21 31 41 51
 MPLEQRSQHC KPEEGLEARG EALGLVGAQA PATEEQQTAS SSSTLVEVTL GEVPAADSPS 60
 PPHSPQGAAS PSTTINYTLW RQSDDESSNQ EEEGPRMFPD LESEFQAAS RKMVELVHFL 120
 LLKYRAREFV TKAEMLESVL RNCQDFFPVI FSKASEYLQL VFGIEVVEVV FISHLYLIVT 180
 CLGLSYDGLL QDNQVMPKIG LLIIVLAIIA IEGDCAPEEK IWEELSMLEV FEGREDSPVA 240
 HPRKLLMQDL VQENYLEYRQ VPGSDPACYE FLWGFPRALIE TSYVKVLHHT LKIGGEPRHS 300
 YPPLHERALR EGEE

Seq ID NO: 207 DNA sequence
Nucleic Acid Accession #: NM_021115
Coding sequence: 743-2893

1 11 21 31 41 51
 AAAGGAAGGG AGGGAGGGAG AAAGGAGAAG TTGGTTTAGA GGCCAGCCGG ACGAGCTTTG 60
 GGCACCGCCC TTAGGAGGGC CACCTCAGA GTCTGACAGC AGGTGAAGGT CCTAAATCTC 120
 CCCAACTAA CTGTGTCTTT TTCTCTCTTT CCAAGATGCT CTTCCCGAGG GAGATGCTAG 180
 CCTTTGGGT CCTTACCTCC TGCCCTCAGG AGCCCGGAG AGAGGCACTC CTGGCAAGA 240
 GCACCTGAA GAGAGAGTGG TAACAGCGCC CCCAGTTCC TCACAGTCCG CGGAAGTGCT 300
 GGGCGAGCTG GTGCTGGATG GGACCGCACC CTCTGCACAT CACGACATCC CAGCCCTGTC 360
 ACCGCTGCTT CCAGAGGAGG CCCGCCCAA GCACGCTTG CCCCCAAGA AGAAACTGCC 420
 TTGCTCAAG CAGTGAACT CTGCCAGGAA GCAGCTGAGG CCCAAGGCCA CCTCCGAGC 480
 CACTGTCCAA AGGCGAGGGT CCCAGCCAGC GTCCAGGGC CTAGATCTCC TCTCCTCCTC 540
 CACGGAGAAG CCTGGCCAC CGGGGACCC GGACCCATC GTGGCCTCCG AGGAGGCATC 600
 AGAAGTGCCC CTTTGGCTGG ACCGAAAGGA GAGTGGCGTC CCTACAACAC CCGCACCCCT 660
 GCAAATCTCC CCCTTCACTT GCGAGCCCTA TGTGGCCAC ACTCTCCCC AGAGGCCAGA 720
 ACCCGGGAG CCTGGCCTG ACATGGCCCA GAGAGCCCCC CAGGAGGACA CCAGCCCAT 780
 GGCCTGTAT GACAAAGGTG AGAATGAGCT GACTGGGTCA GCCTCAGAGG AGAGCCAGGA 840
 GACCACTACC TCCACCATTA TCACCACCAC GGTATCACC ACCGAGCAGG CACCAGCTCT 900
 CTGCAGTGTG AGCTTCTCCA ATCTGAGGG GTACATTGAC TCCAGCGACT ACCCACTGCT 960
 GCCCTCAAC AACTTTCTGG AGTGACATA CAACGTGACA GTCTACACTG GCTATGGGGT 1020
 GGAGCTCCAG GTGAAGAGTG TGAACCTGTC CGATGGGAA CTGCTCTCCA TCCGCGGGGT 1080
 GGACGCGCCT ACCCTGACCG TCTGGCCAA CCAGACACTC CTGGTGGAGG GGCAGGTAAT 1140
 CCGAAGCCCC ACCAACACCA TCTCGTCTA CTTCCGAGC TTCCAGGACG ACGGCTTGG 1200
 GACCTTCCAG CTTCACTACC AGGCCTTCAT GCTGAGCTGC AACTTTCCCC GCGGCTCTGA 1260
 CTCTGGGGAT CTCACGGTGA TGGACCTGCA CTCAGGTGGG GTGGCCCACT TTCAGTGCCA 1320
 CCTGGGCTAT GAGCTCCAGG GCGCTAAGAT GCTGACATGC ATCAATGCCT CCAAGCCGCA 1380
 CTGGAGCAGC CAGGAGCCCA TCTGCTCAGC TCCTTGTGGA GGGGCACTG ACAAATGCCAC 1440
 CATCGGCCGC GTCTCTCTCC CAAGTTACCC TGAACACACA AATGGAGGCC AATTCTGCAT 1500
 CTGGACGATT GAAGCTCCAG AGGCCAGAA GCTGCACCTG CACTTTGAGA GGCTGTGCT 1560

5
10
15
20
25

```

GCATGACAG GACAGGATGA CGGTTACAG CGGGCAGACC AACAGTCAG CTCTTCTCTA 1620
CGACTCCCTT CAAACCGAGA GTGTCCCTTT TGAGGGCCTG CTGAGCGAAG GCAACACCAT 1680
CCGATCGAGG TTCAAGTCCG ACCAGGCGCG GCGCGCCTCC ACCTTCAACA TCCGATTGTA 1740
AGCGTTTGAG AAGGGCCACT GCTATGAGCC CTACATCCAG AATGGGAAGT TCACTACATC 1800
CGACCCGACC TATAACATTG GGAATATAGT GGAGTTCACC TGCGACCCCG GCCACTCCCT 1860
GGAGCAGGCG CCGGCCATCA TCGAATGCAT CAATGTGCGG GACCCATACT GGAATGACAC 1920
AGAGCCCTCG TGCAGAGSCA TGTGTGGTGG GGAGCTCTCT GCTGTGGCTG GGTGGGTATT 1980
GTCCCCAAGC TGGCCCGAGC CCTACGTGGA AGGTGAAGAT TGTATCTGGA AGATCCACGT 2040
GGGAGAAAGG AAACGGGATC TCTTAGATAT CCAGTTCCTG AATCTGAGCA ACAGTGACAT 2100
CTTGACCATC TAGCATGGCG ACGAGGTCTT GCCCCACATC TTGGGGCAGT ACCTTGGGAA 2160
CAGTGGCCCC CAGAAACTGT ACTCCTCCAC GCCAGACTTA ACCATCCAGT TCCATTGGGA 2220
CCCTGCTGGC CTCATCTTTG GAAAGGGCCA GGGATTTATC ATGAAGTACA TAGAGGTATC 2280
AAGGAATGAC TCCTGCTCGG ATTTACCCGA GATCCAGAAT GGCTGGAAAA CCATCTTCTA 2340
CAGCGAGTTG TGCCGGGGAG CCAGAACTAC CTACCAAGTG GACCCCGGCT ATGACATCGT 2400
GGGGAGTGAC ACCCTCACCT GCCAGTGGGA CCTCAGCTGG AGCAGCGACC CCCCATTGTT 2460
TGAGAAAATT ATGTACTGCA CCGACCCCGG AGAGGTGGAT CACTCGACCC GCTTAATTTT 2520
GGATCCTGTG CTGCTGGTGG GGACCCACAT CCAATACACC TGCAACCCCG GTTTTGTGCT 2580
TGAAGGAGAT TCTCTTCTGA CCTGCTACAG CCGTGAAACA GGGACTCCCA TCTGGAGCTC 2640
TCGCTGCCCC TCAGCGGTTT CAGAAGCGCG AGCAGAGACG TCGCTGGAAG GGGGGAAACAT 2700
GGCCCTGGCT ATCTTATACC CGTCTCTCAT CATCTCCTTA CTGCTGGGAG GAGCCTACAT 2760
TTACATCACA AGATGTGCTT ACTATTCCAA CCTCGCCTG CCTCTGATGT ACTCCACCC 2820
CTACAGCCAG ATCAGCGTGG AAACCGAGTT TGACAAACCC ATTTACGAGA CAGGGGGAAC 2880
CCAAAAGGTT TAGGGTTTCA TTTAAAAGA GGTACCCCTT AAAAAGGGGC TTGTGAACCT 2940
AACCCCAATT TCCCCGAGAC ATTTATCCAA AGGCCCTGGG GGCCTTGATT TAAACCCCCA 3000
AAAGCGCGCT GTTTTGTGTT TAAACTTTTT AACAAAGGTT TACGGGTTTT TTCCCCGGAT 3060
TTTATAAATT TTAAGAAGTG

```

30 Seq ID NO: 208 Protein sequence:
Protein Accession #: NP_066938

35
40
45

```

1 11 21 31 41 51
MAQRAPOEDT SPMALMDKGE NELTGSASEE SQETTTSTII TITVITTEQA PALCSVSEFSN 60
PEGYIDSSDY PLLPLNLFLE CTYNVTVYTG YGVELQVKSV NLSDGELLSI RGVGDTLTV 120
LANQTLLEVG QVIRSPNTI SVYFRTPQDD GLGTFQLHYQ AFMLSCNFPF RPDSDGVTVM 180
DLHSGGVAHF HCHLGYELQG AKMLTCTINAS KPHWSSQEPF CSAPCGGAVH NATIGRVLSP 240
SYPTENGSGQ FCIWTIEAPE GQKLHLHFER LLLHDKDRMT VHSQGTNKSA LLYDSLQTES 300
VPEGLLSLSE NTIRIETSD QARAASTFNI RFEAFKGGHC YEPYIQNGNF TTSDFTYNIG 360
TIVEFTCDPG HSLEBQPAII BCINVRDPYV NDTEPLCRAM CGSLSAVAG VVLSFNWPEP 420
YVBGEICIMK IHVGEERKIF LDIQFLNLSN SDIILTYDGD EVMPHILGQY LGNSGPQKLY 480
SSTPDLTIQF HSDPAGLIPF KGQGFIMNYI EVSRNDSKSD LPEIQNGWKT TSHTELVRGA 540
RITYQCDPGY DIVGSDTLTC QMDLSWSSDP PFCEKIMYCT DPEVDHSTR LISDPVLLVG 600
TTIQTCTNPG FLVGGSSLLT CYSRETGTPI WTSRLPHCVS EAAETSLEG GNMALAIPIP 660
VLIISLLLG AIYIYITCRY YSNLRLEPMY SHPYSQITVE TEFDNPIYET GGTQKV

```

50 Seq ID NO: 209 DNA sequence
Nucleic Acid Accession #: NM_001327.1
Coding sequence: 89-631

55
60
65

```

1 11 21 31 41 51
AGCAGGGGGC GCTGTGTGTA COGAGAATAC GAGAATACCT CGTGGGCCCT GACCTTCTCT 60
CTGAGAGCCG GGCAGAGGCT CCGAGGCCAT GCAGGCCGAA GGCCGGGGCA CAGGGGGTTC 120
GACGGGGGAT GCTGATGGCC CAGGAGGCCC TGGCATTCTT GATGGCCAGG GGGGCAATGC 180
TGGCGGCCCA GGAGAGGCGG GTGCCACGGG CGGCAGAGGT CCCCAGGGCG CAGGGGCGAG 240
AAGGGGCTCG GGGCCGGGAG GAGGCGCCCC GCGGGGTCGG CATGGCGGCG CGGCTTCAGG 300
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGGAG AGCCGCTGCG TTGAGTCTTA 360
CCTCGCCATG CCTTTCGCGA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420
GGATGCCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCAGTG TGTCCGGCAA 480
CATACTGACT ATCCGACTGA CTGCTGACAG CCACCGCCAA CTGCACTCTT CCATCAGCTC 540
CTGTCTCCAG CAGCTTTCCC TGTGTATGTG GATCAAGCAG TGCTTTCTCG CCGTGTITTT 600
GGCTCAGCCT CCCTCAGGGC AGAGGCGCTA AGCCAGCCTT GCGCGCCCTT CCTAGGTGAT 660
GCCTCTCTCC CTAGGGAATG GTCCAGCAGC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720
GTTTGTGCTT GAGGAGGAC GGCCTACATG TTTGTTCTG TAGAAAATAA AACTGAGCTA

```

70 Seq ID NO: 210 Protein sequence:
Protein Accession #: NP_001318.1

75

```

1 11 21 31 41 51
MQAEGRTGG STGDADGPGG PGIPDGPQGN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60
PRGPHGGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EAELARRSLA QDAPPLVPVG 120
VLLKEFTVSG NILTIRLTAA DHRQLQLSIS SCLQQLSLLM WITQCFLPVF LAQPPSQRR

```

80 Seq ID NO: 211 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 52-459

85

```

1 11 21 31 41 51
CCTCGTGGGC CCTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGGAGC CATGCAAGGC 60
GAAGGCGCAG GCACAGGGGG TTGACAGGGC GATGCTGATG GCCCAGGAGG CCCTGGCATT 120
CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CCGGTGCCAC GGGCGGCAGA 180
GGTCCCGCGG GCGCAGGGGC AGCAAGGGCC TCGGGGCGGA GAGGAGGCGC CCGCGCGGGT 240
CCGCATGGCG GTGCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTGCGGGGC CAGGAGGGCG 300

```

5 GACAGCGGCC TGCCTCAGTT CCGACTGACT GCTGCGAGCC ACGGCCAACT GCAGCTCTCC 360
ATCAGCTCCT GTCTCCAGCA GCTTTCCTGT TTGATGTGGA TCACGCAGTG CTTTCTGCCC 420
GTGTTTTTGG CTCAGGCTCC CTCAGGGCAG AGGCGCTAAG CCCAGCTTGG OGCCTCTTCC 480
TAGGTCTATG CTCTCTCCCT AGGGAATGGT CCCAGCAOGA GTGGCCAGTT CATTGTGGGG 540
GCCTGATTGT TTGTGCTGG AGGAGGACGG CTTACATGTT TGTTCCTGTA GAAATAAAG 600
CTGAGCTA

Seq ID NO: 212 Protein sequence:
Protein Accession #: Eos Sequence

10 1 11 21 31 41 51
MQAEGQGTGG STGDADGPGG FGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60
15 PRGPHGGAAS AQDGRCPGCA RRPDSRLLOF RLTAADHRQL QLSISSCLQQ LSLLMWITQC 120
FLPVFLAQAP SQQRR

Seq ID NO: 213 DNA sequence
Nucleic Acid Accession #: NM_000555
Coding sequence: 416..1498

20 1 11 21 31 41 51
CTTATTTTTT ATGAATGTGG GATAGCTGCA CCAGCTTGGT GGGGAAAGGG TTGATGAAT 60
25 AGCACAAAGA CACTGGCTGT TCCCTGGAGG CTGTCCCTTT AAAGGAGAAT CTTAGTTTAT 120
TCTGGGGGGA GGGGATGCAC ACATTAGAGT AGGAAAGAGG GCTTGAATAA AAATGAAAC 180
ACTCCCCCTT CATAGTCATT GTACTGAAAT GCAGGAGCTG CTTCTTAAGC TGGAGATGCT 240
AACCTTGGGT AGCTCCTTCT GTTCTCTTCA AGGGGAATTT TGTGAGGCTA TGGATTCTAT 300
TACAACTGTT AGTCATGTGG GCATGTGTGA GGAAACAGAT GCCAGTTTAA ATGTATTTAG 360
CCCGAAGTTC CAAATTGTATA GGAGCCACTG TCAGTCTCTG AGTTTCCACC AAAATATGGA 420
30 ACTTGATTTT GGACACTTTT ACGAAAGAGA TAAGACATCC AGGAACATGC GAGGCTCCCG 480
GATGAATGGG TTGCTTAGCC CCACTCAGAG CGCCCACTGT AGCTTCTACC GAACACAGAA 540
CTTGCAAGCA TGGTCTGTCA AGAAGAAAGC CAAGAAGGTA CGTTTCTACC GCAATGGGGA 600
CCGCTACTTC AAGGGGATTG TGTACGCTGT GTCTCTGTAC CGTTTTCGCA GCTTTGACGC 660
CTTGTGCTGG GACCTGACGC GATCTCTGTC TGACACATC AACCTGCTCT AGGAGTGGCG 720
35 TTACATTTAC ACCATTGATG GATCCAGGAA GATCCGGAAG ATGGATGAAC TGGAGGAAGG 780
GGAAAGCTAT GTCTGTCTCT CAGACAACTT CTTTAAAGAG GTGGAGTACA CCAAGAAATG 840
CAATCCCAAC TGGTCTGTCA ACGTAAAGAC ATCTGCCAAT ATGAAAGCCC CCCAGTCTCT 900
GGCTAGCAGC AACAGTGCAC AGGCCAGGGA GAACAAGGAC TTGTGCGGCC CCAAGCTGGT 960
TACCATCATC CCGAGTGGGG TGAAGCCTCG GAAGGCTGTG CGTGTGCTTC TGAACAAGAA 1020
40 GACAGCCAC TCTTTTGGAG AAGTCTTCAC TGATATCACA GAAGCCATCA AACTGGAGAG 1080
CGGGGTGTGC AAAAATCTCT ACACCTCTGA TGGAAACAG GTAACCTGTC TCCATGATTT 1140
CTTTGGTGAT GATGATGTGT TTATTGCTGT TGGTCTGTAA AAATTTGCTC ATGCTCAGGA 1200
TGATTTTCT CTGATGAAA ATGAATGCCG AGTCATGAAG GGAAACCCAT CAGCCACAGC 1260
45 TGGCCCAAG GCATCCCAAC CACCTCAGAA GACTTCAGCC AAGAGCCCTG GTCCTATGG 1320
CCGAAGCAAG TCTTTTGGAG ACTCAGCAAA CGGAACCTCC AGCAGCCAGC TCTCTACCCC 1380
CAAGTCTAAG CAGTCTCCCA TCTCTACGCC CACCACTCCT GGCAGCCCTC GGAAGCACA 1440
GGACCTGTAT TGGCTCTGTG CCTTGGATGA CTCGGACTCG CTTGGTGATT CCATGTAAAG 1500
GAGGGGAGAG TGCTCAGAGT CCAGAGTACA AATCCAAGCC TATCATTGTA GTAGGGTACT 1560
50 TCTGCTCAAG TGTCCAACAG GGCTATTGGT GCTTCAAGT TTTTATTTTG TTGTTGTGT 1620
TATTTTGAAA AACACATTGT AATATGTTGG GTTATTTTTC CTGTGATTTC TCTCTGGGC 1680
CACTGATCCA CAGTTACCAA TTATGAGAGA TAGATTGATA ACCATCTCTT GGGGAGCAT 1740
TCCAGGATGA CAAATGCTGT TAGTCCATGA CCTTTCAATG GAAAGCTTAG GGGCTGGGG 1800
TAAATTTGCC CGTTTTAAAT TTGCCCAAAC AGTTTTCTCT TTGTAGAGGG GTGTTTAAAT 1860
55 ATACAGCAAT TAAAGAGTTT GTGTGGGGAA AAAAAAATC CATTGGCAGA TCCAAGAAAT 1920
ACAAACACAA GTGCCCTTT TCTCTGGATC TCAAGAAATG TGGAGGACCC TGGAAAGACA 1980
GCAAGGCAGC TCCCGACCT CACTCTTAC TCTCTGATGA GGCCCGGGTT TGTGTGCCAG 2040
CACCAATGCT GGCTGCTGAT GGGGAGAAAT AAACCAACAA CTTATAATTG TGACCCAGA 2100
TGCTTAGGAT CCTGTGCTGT GGTTAGCTAA GAGAATAGAC AGAATTGGAA AATACTGCAG 2160
60 ACATTTCCGA AGAGTTTATA AAGCACAGTG AATCTCTGGT CAATCTCTCC ACTGAGGCAA 2220
TTTGGAAATCA ATAGCAATTG GATAATAGTT TGGAGTAAGG GACTTCATAT ACCTGATTCC 2280
TCTAGAGAGC TGTCTAATAT ACCACATGAT TACATGAAT GTATGGTATC CATCTATCTC 2340
TGTTCTATTG AATGCTTGT TAACAGCCAA CACTGAAAC ACTGTGAGAA TTTGTTTCA 2400
GGTCTGACAC CTTTCACTCT CTTTTATAG CAAGAAATCA ATATCTTTT TATAAAAT 2460
65 CATGTCGTA TTTCAAGAGC AAATCTTTCA GGCTCCTTT TTATAAATG GTGATTTTTC 2520
TTTTGTCTAA AAAACACATG AAGAAATTT ACCAGAAAA AAAAAAAG CCGAAGAATA 2580
ATGTTATTTA GAAATATGCT TGTCACTGCC AAACAGTAAC CTCCAGGAGA AAACAGATG 2640
AATAGCAGAG GCCAATTCAA TAGAATCAGT TTTTGATAG CTTTTAACA GTTATGCTTG 2700
CATTAAATAT TTCAATGTGG ACCAGACATT CTAATATAT TTTAAATGAA ATGTTACAGC 2760
70 ATATTTTAAG CAATCTTTT TATCTATAAT CCTAATATT CATCTGAAG ACACAGAAAT 2820
CTTTCACCTG TCTTTAATAT TAGAAAGGAT TTCTCTTAC TAAGGACTGA TCATTGAAA 2880
TAGTTTTCAG TCTTTGAGA TACAGGTTTA TAACACTGCT TTTTTTTTCC TGTAAACATA 2940
GCCATAATG GCAAAACAA CTAATTTTAA TTGAAGGTCT TGCTTGCCAN TCTGTGTTG 3000
GCTTTNACCA AATATAAAAA TTCCTTATT CCTTGGTAAT GGTGCAAAIN TTTGGAAGG 3060
75 CACAGCATCC AAACCAAGCT GCTGTTGGC TACTGAATGG CTGCGAGTTG TTCCTCCACT 3120
CTAAATGGAA TGAGCTTGCT GTGTGTGTGT GTGGTGGTGG TGGGAGGGGG TGGTGCACTG 3180
GTGTGTGTGT GTGTGCACT GCAGCTGCTT CAAATTAAG AAATACTACA AGACACCCCT 3240
GTAATGGATT GGTGGCACT GGTGGTGTGC ACTGTGTAGG GGGGAACCCA 3300
GTGGTGGTGG GGTATCTCAA ATGCCCTAG ACAAGCTTCA GATGTCTGTA GCTACCAAAA 3360
80 ACATTTTCGG TTCAAGAAAA GTGAGATGAT GGTAGTACTG GTTCTGCTG AAATGAAAA 3420
ACCCCAATG ATGAGATCT CTTTTGCCC CCTCTCCTT TTTTGTAAAC CCAATCAAAA 3480
CCATTAAATA GCCCATTTTA CTAANCCCT ATTCTTCT AGAAGCTCAG GGTTCNTTA 3540
GTGCCCTCCA NAACATTTT TAGTTAATTG GGAAGAGTG ATACTTGGAT TAGGGGGTGT 3600
GGGCATAAAG AATGGTGGGA GGCCTGATTT TAAATTCAG GCCAGAACCC CCAATGACTC 3660
CACCATATG NTCATTTAG GTCTCATTTA GTCCATCACC TTTATTTTAA GTTAGGGAAG 3720
85 TGGAGGCTGG TAAAGAGCAG GACCAGAGGA AGAATCCAGA TTTCTTATG CTTGGGCTC 3780
ACACTAGCTC TMTGATATT TCCTTGATTG CGGTATATGT ACTACTAGAA AATACCAAT 3840
GGATATATTT TCTTTAGGAT AACCTTTGAA CCAACAAAT TCAATAACAA TAGTACATCT 3900

	TCCATCTTAC	TTTTAATCGA	GTATAAGGAA	ATGTTTCTTT	ATGGCCATTT	TGGAGGGAGC	3960
	AGGGGATGAG	GCTTGGGCATA	GTCCAAAATT	TAAGNCTCCA	ATAATTAATT	GCATTTTAAA	4020
	TTGTTTTAAA	TTGGCCCACT	ITCAAGGCAA	TTTTTTTTGT	GTGCTGTGAA	CTGAGCTCCT	4080
5	CCACCCCTGT	CATTCACTTC	CAATTTTACC	CAATCCAATT	TTAGCACTCA	AGTTCCATTG	4140
	TGTTAATTTT	TGCACGGTCT	ACACACATCA	AGTCAGCAAG	CATTTGCCAC	CACTCCCTAT	4200
	ACTTCTCCCT	CTTTTTCACA	CACACACACA	CACACACACA	CACAAATCCAT	CTCTTGCTTG	4260
	TTCCCTACCT	CTGATTTTTC	CTTCCCTACA	GAATATAGAA	TAGGGACAAA	GAAGGGGAAA	4320
	ATGTATATAT	TGGGGCTGGG	CTGAACAACT	AACTTCATAA	GTAGTATTAA	CTAGGGGTAA	4380
10	ATTGAGAGAA	AAGCTCCTTT	TCTCTTCACT	GTTTGGGAAA	GGATAGCCAT	TAGCATGACT	4440
	GCTTTGTGTC	CTTATGGACT	TTAGTATTAG	CCTAGATTGA	ATTATAGCGT	TTTTCTAGCT	4500
	GAAGGAACCT	TAAAGTACCA	TCATCTACTC	CTCTACTCCA	AATTTCTCAT	TCTTCAGGCC	4560
	AGGAAACCGA	GACACAGAGG	TAAAGTAATT	TCCCAAGGT	CACACAGCTG	GCTGGGGCAG	4620
	GATTGGGTTT	ACAACCCACA	TCTCCTGGCT	CTTATTCAG	GGCCTTTTCC	CACTAAGTAG	4680
	TATTGCGCTC	CATTAGGCTC	CTGAGAGTTA	TTTCTCAGGG	TCATGTTGCA	TCTTGAGGCC	4740
15	ACATGCTGCT	GCCCTGATCT	CAGTGGGAAA	TNCAACCCAGC	AACCTAATAC	AGCCCCCTTT	4800
	CCCTGCATTG	ACCTGGTTCC	CATCCACATG	GGTTCAGAT	GTCCCTGAAG	AGAGTGAGGC	4860
	ATTGAGGGCC	AATAGGAGCA	ATGGGGTCCC	TGGCCTTGTC	CATCTGATTC	AGGAGATCAC	4920
	TGCTCCATCG	TGAGGAGGCC	TCTGAATAGC	CCCCCACTGA	ATGCTTGCTT	TGCCCAAATG	4980
20	GAATGGAGGA	AGATTGATTT	TCTCCATCAG	TTCACTTGT	GTCTCTCAT	AATGGTTGGT	5040
	CTTTCAGAGC	TGAGGGAAAT	GTTTCTTGTT	TCCANAGTAN	AAAAAGAGAA	GAGTGGAAAC	5100
	ATANCTTTGT	TCATCTTAAC	TTTCTGAGAT	GGCTTTTCAA	CATTTAAAAA	AAACTAGTGT	5160
	GGTACCATTC	ACTGGCANGA	TTTNTTTTAT	AATATGGGAG	TAGATGAGG	TAGAGAAAAA	5220
	AACCTGGTCT	CATCTGGGTT	GCCCTCATCC	ACAATGTCCC	CAAGGCCATC	CTGCTNTGAT	5280
25	GAGGACAAAT	TCCAGGTATA	AGCAAGGGGC	TTTGTGACAA	AAATGTACCC	TGGCTGATGT	5340
	TAAACATTTG	CTCTGTGTTT	TGCACCAAAA	TAGCAAGCTG	TGTGCTCTAT	ACACTCTTCC	5400
	CATCGTCTTG	TGTACATCGC	TCTGTGGGCC	TTCCACAGCA	GAACCCAGGG	CAAAAGGGTC	5460
	CAAAACATCG	GTTTTCTCTG	CTGCAAGGCT	NTTCTGGGGA	ACTAAGGGGG	TATTTATTAG	5520
	TTCACTTNTA	AGAGACCTCC	TTCTGGGCTT	ACCCCACTCC	TCAGTACTT	CTCTCTCCTT	5580
30	CCTCCTTCTC	CTCCACAGTC	ACAAGTAACC	AAGGAACCTG	AAAGTGGATG	TGTAGCTATT	5640
	TGAAGAGGCG	AAGGAACCTG	GAGATTCTTC	TTTGAATCCT	TTAGTCCAAAG	TCTTAGACCA	5700
	GTGATTGGTG	CTTACCTTGA	ACAAAATTTT	GTCTGTGTTT	CTAATCCCTT	CAATACNTG	5760
	GGTACAAATG	TCCCAATCAC	CCTGCACATT	TGATTCTAAA	TGGCTTTTAT	TTTTTAAAAA	5820
	TCCATATCCC	TAGGACAGAA	NAACAGGATG	CCTATATCCC	CAAAATGAGC	TCCAGGACAC	5880
35	TGATGGGAAT	GATCCCAANG	ATCACCACAC	CTCAGAAAAC	GTCTGTGCCA	ANAGACTTCC	5940
	CCAGATAGAA	NCACCTGGGAC	AGTGGTTTGA	ACGACTCTCT	TTATGGTTGT	CCAGTTTGCT	6000
	ATGGAAATAA	AAGGCATTTA	TTTTTAAAAA	AAGATGATGG	GAACCTGTCT	TTGGCCACAT	6060
	AGGGCCATCT	GGATCCATTT	CCAGGCCCTA	CTCATATATT	GCCTTCACTG	AAGGGCTTTG	6120
	GCTTTAAGTC	CCAGACTGGT	CTCCCAAGTG	AACCATAAGT	GTTTTGGAGC	TCATCTGGGG	6180
40	TGAGGCATGA	GAATGTTGCC	CCATCTATCC	CTTCAGGAAA	AGGTGCCTTC	CCTCCCTTTC	6240
	TCCTAAAGCT	TGGTCCCCAA	AAATGTTTTT	TGCTCCAAA	AGCTAGTAT	GGTCTTTATA	6300
	CACCCANACT	CTTAGTGTGG	CGTCTGCCTT	TGTTTCCTTG	TTAAGGATCT	ATGCANACCT	6360
	CCCGCTTTGG	CTTAGCTTGG	GTGACATTGG	CTATCAITTG	ACAAGACTAA	CTTTTTTTTT	6420
	TTTTTTTTTG	ACTGAGTCTC	CCTCTGTCAC	CTAGGCTGGA	GTGCACTGGC	ACAATCTTGG	6480
45	CTGCTGCAAA	CCCTTACCCCT	TCACCTCCCA	GGTGGAAAGG	ATTCTCCCTG	CTCAGTCTCC	6540
	CGAGTAGCTG	GGATTACAGG	CGTGGGCCAC	CAAACTCTGG	TATTTTTTTA	TTAATTATTAT	6600
	TTTTAGTAGA	GATGGGGTTT	CACCATGTTG	GCCAGACTGG	TCTTGAATCT	TTGGCCTCAA	6660
	ATTATCTGCC	CACCTCGGCC	TCCCAAAGTG	CTGGGATTAC	AGGCATGAGC	ACCATGCCCA	6720
	GCTGACAGAA	CTAATTTTTT	ATCCCTTGGT	TTATTGGCTT	CAACATCTTC	TGGAATCAGA	6780
50	GGTGATTTTT	TCTTACCTTG	GATGCTGTAG	ACTAGGGGAG	TATAGAATTC	CAATTTGGTA	6840
	TTAAGGCATC	TTCTGCTCC	TGATCAGAA	GGCAGGTTAG	TTGGGAGAGG	TCAGATGGCA	6900
	CAACAGAAAT	CACCTGTGAA	GTAAGGCAAA	GACTTTGAAG	GCATTAGCGT	TTCTCATTTAC	6960
	TTAGGTCAAT	AACCTTGAGG	GAATCAATGG	CTTTTGTGCC	GCTCTACCTC	TTTGTGTATC	7020
	TCCTTGACTT	TTCTTCTCTT	GTCTAGTTTC	CTCTGTCTCT	AGTTTATATT	CTATGTTATC	7080
55	AGTCTCTCTT	TCCACAGTAC	AAACATCCAT	CCTTCTCTCT	GTGCAATTC	GTCTCTCCCT	7140
	CTTATTATCT	TTTCTGCTAC	TTTCTCTCTC	CTCCCTGTCT	AGGCATTTGG	CATGTGCCCTC	7200
	TTCTTAGCCT	GTGATTTTGC	CTTGGGACTG	ATGATAAATT	ATTCCAGAT	TCAATCAGCC	7260
	CTGCTCCTAC	CCCAGTCCAA	TCAGAAGTAT	GTGGTGGGG	AATCAACCTG	ATCCTGGCCC	7320
60	TTCTTCTTTC	TTCATTTTCA	TTGGTAATCC	CCCTCAGCAG	ATCTTTACAA	GCAGTTTCTT	7380
	TATAGCTCAT	GTACTTTTAG	GTCTTTCCTT	TCCAAGCACT	GTACAGAAAT	CTTTGTGGTT	7440
	CCTTTTTAGT	CTGACATTTT	GTGGAGCAGT	GAAGCGTGCT	CAGAGACATA	ATCAGCTGAA	7500
	GAGAAAAAAT	CCACCCATGG	ATTATATATCA	GCTAAATACT	AATAATTGAT	TTTGTTTGAT	7560
	GTGCCCATAA	TTTTTAAAGC	TGCAATATAA	TATAATGAGG	GACCAAGGT	AATTTCTCCT	7620
	GTCAATTTGT	TGCGCTGGAT	GGGGGTGGGG	GAGTAATTGC	TTAAAGTTTT	ACCATTACAC	7680
65	ATTAAACTCT	CTATAATAAT	CTTGTTTGGG	GCTTGCTAAC	TGTTGAGCTG	TTTTAACTAA	7740
	ACTGGTAGGC	AATCGGAGTT	GATTTAAATG	AAAAGATAAT	TTAACAATC	TATACTATAA	7800
	AAAGAGACAT	TGCTTAATTT	GACATGTATT	TTTTCTTCT	GAGTCACCTA	AACATTTACT	7860
	CTTGACACCA	ACTGTTTCAT	ATACTGAATA	GACAGTCCAT	ATAAGAGAAA	TTAGTGGACC	7920
70	TAAAGAAGCC	AGATTGTAGG	TGTTAATTTA	TTAAACAGAA	TTGCAAGGCC	CTTGGAAATG	7980
	TCACCTGCTT	GCAATACCAT	ATGGCATGCC	AAAATTTACA	ATGACTTTTC	TTTATAAGTT	8040
	ATCCAAAGG	GATTTGAACA	AGTAAGAGGT	TATGCCAAAA	TGCTCCCAAT	GTATGGTCTT	8100
	GTAATATATT	GCAGCTTGAA	GCCATATGAT	CCTATGACT	TGTATACAAC	TAAATGATGT	8160
	TTTATTGAAT	TTTGCAATTC	CCAGTGTGG	TAAAGCTTTA	AAATGTTTTT	GATCACCTTT	8220
	MTGTGCCATT	AAACTTGTAT	AGAAAAATGT	TTTATGGCCA	TTTTCAAGG	GAGAAAGTTT	8280
75	AAAAATGAAA	CAGCCACCCC	TTTCTGCCCT	ATAGCTGTAG	TTAGAATTTA	GTACCTGTAG	8340
	CAAAACAGCT	GTAATTGGTG	GTTGTAGTGT	TAGAGGTGTT	AGCTTGCTAG	TGACTAGCTT	8400
	TGGAGAGTAA	ATGCATGSTA	TTGTACATCA	CATTTCITAA	CTGTTTTTAA	CCTCTGAAAA	8460
	GAATATATTC	TTCTTTGTAG	TCCTTCTTCC	CACCCCTTGT	CCCTCTCCCT	CTCCCTGCTC	8520
	CCAGTTGTCT	TACAGTTTGA	AATATCTGAT	TTGAGGCCCA	ATAACTCTTG	CCAAGTAAAG	8580
	TCAGCAACCA	ACAAACAAAC	CAAAATGTGG	GGAAAAAGCA	TTTCTCAACC	ATCTCTCAGC	8640
80	AGTTATTGAT	CATTTCTTAA	GGAAACAGCAT	TGTGATCAAA	GACTCAACTT	TACGTAAAAA	8700
	TCAGTGGTAA	ATTGGGGTTG	TATTGGCCAT	TGATTACATT	CAGGATTGAA	TAGTTTTTCA	8760
	AATCAGATGT	AATCCAAAGA	CAGTAGGTAG	TGATGTCCTT	TATCCCTGCA	GCTGTTTTAA	8820
	GATAGAGACC	TCAGAGAGAT	CTGCTTGACC	GATGACCAAT	AATTATTGTA	AAAAAAAAGA	8880
	AAAAATGAGA	GAATAAAAAC	AGATATTATA	GAACCTTTAGC	CACCTATTTA	GAATAGTTAT	8940
85	AGCCAGAAAA	AAAAACAAGG	GCAATGAGTT	AAATGCATTA	CTATCAGTGT	CCTAGGCAAT	9000
	ACCTAACCTA	CTCTGAAATT	GTGATTCAAA	AGCAGTATTT	CAAGAGGCAT	TCTCCTTTTT	9060
	TGGTTTGTCT	ACCCCACTTG	GACTGGTAGG	TTTGGTGAGG	CCCCCATAAA	CCAGCTGGAG	9120

WO 02/086443

PCT/US02/12476

CAGACCTTTT TCATCTCTG TGCTGTAAAC ACCCTCTCTC CCCCACCCCC TCOCGAATTC 9180
AATGAGGGCT TCTTTGGGTC AGAGGACTTC AAGGTTGTCT AGAGAAGTTT GCCATGTGTG 9240
TAAGTGCTG TGAACTGTGA GTGCTGAAGA TTGCGAGCAT TCAATACCAG GCAGCCAAAG 9300
AGCTGCTCTT GCAATTATTT TGGCTCTCAA GCTCTGTTCT TCATCGCATT CTCATTTCTG 9360
TGTACATTTG CAAGATGTGT GTAATGTCTAT TTTCCAAAAA TAAATTTGA TTTCAAT

Seq ID NO: 214 Protein sequence:
Protein Accession #: NP_000546

1 11 21 31 41 51
MELDPGHFDE RDKTSRNMRG SRMNLPSPT HSAHCSFYRT RTLQALSNEK KAKKVRFYRN 60
GDRYFKGIVY AVSSDRFRSF DALLADLTRS LSDNINLPQG VRYIYTIDGS RKIGSMDELE 120
EGESYVCSSE NFFKKVEYTK NVNPNWSVNV KTSANMKAPQ SLASSNSAQA RENKDFVRPK 180
LVTTIRSGVK PRKAVRVLNL KKTASFEQV LTDITEAIKL ETGVVKKLYT LDGKQVTCILH 240
DFFGDDDVFT ACGPEKFRYA QDDFSLDENE CRVMKGNPSA TAGPKASPTP QKTSKSPGP 300
MRRSKSPADS ANGTSSSQLS TPKSKQSPIS TPTSPGSLRK HKDLYLPLSL DSDSLGDSM

Seq ID NO: 215 DNA sequence
Nucleic Acid Accession #: NM_130467
Coding sequence: 312..644

1 11 21 31 41 51
GGCAGGAGGC AGAGCTCTGC AAGGAGAGGT TGTGTCTTGG TTCTTTCCGC CATCTTCTGT 60
CTTTCCAAACA TCTTCGTTCT TTCTCACTGA CCGAGACTCA GCCGTAGGT CTGCAGAGTG 120
GTCTTCTCTG TAAATTAAGT GTGAGTGAAT GTGTGGAGGA GCCAGCGGGC TTAGGACAGG 180
TCTGTGCGCA CAGTCCGTGG CTTTGAGGGA AAAGGCCCTC GCGGTGGTCC TCCGCTCTCC 240
CCCAGGTCGT GATGACGGCG CCATGGGCGG GTAATCGTGG CTGGGCTGGA ACGAGGAGAG 300
AAGTGAGAGA TATGAGTGAG CATGTAACAA GATCCCAATC CTCAGAAAGA GGAAATGACC 360
AAGAGTCTTC CAGGCCAGTT GGACCTGTGA TTGTCCAGCA GCCACTGAG GAAAAACGTC 420
AAGAAGAGGA ACCCAAACT GATAATCAGG GTATTGCACC TAGTGGGGAG ATCAAAAATG 480
AAGGAGCACC TGCTGTTCAA GGGACTGATG TGGAGCTTT TCAACAGGAA CTGGCTCTGC 540
TTAAGATAGA GGATGCACCT GGAGATGGTC CTGATGTCAG GGAGGGGACT CTGCCCACTT 600
TTGATCCAC TAAAGTCTG GAAGCAGGTG AAGGGCAACT ATAGGTTTAA ACCAAGACAA 660
ATGAAGACTG AAACCAAGAA TATGTCTCTT ATGCTGGAAA TTGACTGCT AACATCTCT 720
TAATAAAGTT TTACAGTTTT CTGCAAAAAA AAAAAA AAA

Seq ID NO: 216 Protein sequence:
Protein Accession #: NP_569734

1 11 21 31 41 51
MSEHVTRSQS SERGNDQESS QPVQFVIVQ PTEKRQEEB PPTDNQGIAP SGEIKNEGAP 60
AVQGTDEAF QQELALLKIE DAPGDGPDVR EGTLPFTDPT KVLEAGEGQL

Seq ID NO: 217 DNA sequence
Nucleic Acid Accession #: NM_001476.1
Coding sequence: 82..435

1 11 21 31 41 51
GCCAGGAGC TGTGAGGAGC TGCTGTGTGG TTCTGCGGT CCGGACTCTT TTCTCTCTAC 60
TGAGATTCAT CTGTGTGAAA TATGAGTTGG CGAGGAAGAT CGACCTATTA TTGGCCTAGA 120
CCAAGGCGCT ATGTACAGCC TCCTGAAGTG ATTGGGCTTA TGGGCGCGA GCAGTTCAGT 180
GATGAAGTGG AACCAGCAAC ACCTGAAGAA GGGGAACAG CAACTCAACG TCAGATCCT 240
GCAGCTGCTC AGGAGGGAGA GGATGAGGGA GCATCTGCAG GTCAAGGGCC GAAGCCTGAA 300
GCTGATAGCC AGGAACAGG TCACCCACAG ACTGGGTGTG AGTGTGAAGA TGGTCTGAT 360
GGGACGAGG TGGACCGGCC AAATCCAGAG GAGGTGAAAA CGCCTGAAGA AGGTGAAAG 420
CAATCAGAG GTTAAAGAA GACACGTTGA AATGATGCAG GCTGCTCCTA TGTGGAAAT 480
TTGTTCAATTA AAATCTCTCC AATAAAGCTT TACAGCCTTC TGCAAAA

Seq ID NO: 218 Protein sequence:
Protein Accession #: NP_001467.1

1 11 21 31 41 51
MSWRGRSTYY WPRPRRYVQF PEVIGPMRPE QPSDEVEPAT PEEGEFATQR QDPAAAEQGE 60
DEGASAGQGP KPEADSQEQG HPQTGCECED GPDGQEVDPF NPEEVKTPEE GEKQSQC

Seq ID NO: 219 DNA sequence
Nucleic Acid Accession #: NM_001476
Coding sequence: 90-3671

1 11 21 31 41 51
ACAGCGGAGC GCAGAGTGAG AACCAACCAAC CGAGGCGCGG GGCAGCGACC CTGCAGCGG 60
AGACAGAGAC TGACGCGGCC GGCACCGCCA TGCCCTGCGCT CTGGCTGGGC TGCTGCCTCT 120
GCTTCTCGCT CTTCTCTGCC GCAGCCCGGG CCACCTCCAG GAGGGAAGTC TGTGATTGCA 180
ATGGGAAGTC CAGGCAAGTGT ATCTTTGATC GGGAACCTCA CAGACAAACT GGTAAATGGAT 240
TCCGCTGCGT CAACTGCAAT GACAACACTG ATGGCATTCA CTGCGAGAAG TGCAAGAAATG 300
GCITTTACCG GCACAGAGAA AGGACCGCCT GTTTGCCCTG CAATTGTAAC TCCAAGGTT 360

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CTCTTAGTGC TCGATGTGAC AACTCTGGAC GGTGAGCTG TAAACCAGGT GTGACAGGAG 420
CCAGATGGGA CCGATGTCTG CAGGCTTCC ACATGCTCAC GGATGCGGGG TGCACCCCAAG 480
ACCAGAGACT GCTGAGACTCC AAGTGTGACT GTGACCCAGC TGGCATCGCA GGGCCCTGTG 540
ACGCGGGCCG CTGTGTCTGC AAGCCAGCTG TTAAGTGGAG ACGCTGTGAT AGGTGTGCGT 600
CAGGTTACTA TAATCTGGAT GGGGGGAACC CTGAGGGCTG TACCCAGTGT TTCTGCTATG 660
GGCATTTCAG CAGCTGCCGC AGCTCTGCAG AATACAGTGT CCATAAGATC ACCTCTACCT 720
TTTATCAAGA TGTGTATGGC TGGAAAGGCTG TCCAAACGAA TGGGTCTCCT GCAAAGCTCC 780
AATGGTCACA GCGCCATCAA GATGTGTTA GCTCAGCCCA ACGACTAGAC CCTGTCTATT 840
TTGTGGCTCC TGCCAAATTT CTGGGAATC AACAGGTGAG CTATGGGCAA AGCCTGTCCT 900
TTGACTACCG TGTGGACAGA GGAGGAGAGC ACCCATCTGC CCATGATGTC ATTCTGGAAG 960
GTGCTGTGCT ACGATTCACA GCTCCCTTGA TGCCACTTGG CAAGACACTG CCTTGTGGGC 1020
TCACCAAGAC TTACACATTC AGGTAAATG AGCATCCAAG CAATAATTGS AGCCCCAGC 1080
TGAGTTACTT TGAGTATCGA AGGTACTGCG GGAATCTCAC AGCCTCCGC ATCCGAGCTA 1140
CATATGGAGA ATACAGTACT GGGTACATTG ACAATGTGAC CCTGATTTC GCGCCGCTG 1200
TCTCTGGAGC CCGAGCACCC TGGGTTGAAC AGTGTATATG TCCTGTTGGG TACAAGGGGC 1260
AATTCTGGCA GGATTTGTCT TCTGGCTACA AGAGAGATTG AGCGAGACTG GGGCCTTTG 1320
GCACCTGAT TCTGTGTAAC TGTCAAGGGG GAGGGGCTG TGAATCCAGC ACAGGAGATT 1380
GTTATTTCAG GGATGAGAAT CCTGACATTG AGTGTGCTGA CTGCCAATT GGTCTTACA 1440
ACGATCCGCA CGACCCCCGC AGCTGCAAGC CATGTCCCTG TCATAAGCGG TFCAGCTGCT 1500
CAGTATGCC GGAGACGGAG GAGTGGTGT GCAATAACTG CCTCCCGGG GTCCACGGTG 1560
CCGCTGTGTA GCTCTGTGCT GATGGCTACT TTGGGGACCC CTTTGGTGAA CATGCCCCAG 1620
TGAGGCTTTG TCAAGCCTGT CAATGCAACA ACAATGTGGA CCCCAGTGCC TCTGGGAATT 1680
GTGACCGGCT GACAGGCAGG TGTGTTAAGT GTATCCACAA CACAGCCGCG ATCTACTGCG 1740
ACCAGTGCAG AGCAGGCTAC TTGGGGGACC CATTTGGCTCC CAACCCAGCA GACAAGTGTC 1800
GAGCTTGCAA CTGTAAACCC ATGGGCTCAG AGCCTGTAGG ATGTGGAAGT GATGGCACCT 1860
GTGTTTGCAA GCGAGGATTG GGTGGCCCCA ACTGTGAGCA TGGAGCATTC AGCTGTCCAG 1920
CTGTCTATAA TCAAGTGAAG ATTCAATGG ATCAGTTTAT GCAGCAGCTT CAGAGAATGG 1980
AGGCCCTGAT TTCAAAGGCT CAGGCTGGTG ATGGAGTAGT ACCTGATACA GAGCTGGAAG 2040
GCAGGATGCA GCAGGCTGAG CAGGCCCTTC AGGACATTCT GAGAGATGCC CAGATTTTCA 2100
AAGGTGCTAG CAGATCCCTT GGTCTCCAGT TGGCCAAAGT GAGGAGCCAA GAGAACAGCT 2160
ACCAGAGCCG CCTGGATGAG CTCAAGATGA CTGTGGAAG AGTTCGGGCT CTGGGAAGTC 2220
AGTACCAGAA CCGAGTTCCG GATACTCACA GGCTCATCAC TCAGATGCAG CTGAGCCTGG 2280
CAGAAAGTGA AGCTTCTCTG GGAACACTA ACATTCTGCG CTCAGACCAC TACGTGGGGC 2340
CAATGGCTT TAAAGTCTG GCTCAGGAGG CCAACAAGAT AGCAGAAAGC CAGCTTGAGT 2400
CAGCCAGTAA CATGGAGCAA CTGACAAAGG AAACCTGAGG CTATTCCAAA CAAGCCCTCT 2460
CACTGTGCG CAAGGCCCTG CATGAAGGAG TCGGAAGCGG AAGCGGTAGC CCGGACGGTG 2520
CTGTGTGCA AGGGCTGTG GAAAAATTG GAAAAACCAA GTCCCTGGCC CAGCAGTTGA 2580
CAAGGGAGGC CACTCAAGCG GAAATTGAAG CAGATAGGTC TTATCAGCAC AGTCTCCGCC 2640
TCTCTGATTC AGTGTCTCGG CTTCAGGAG TCAGTGTATC GTCCCTTCAG GTGGGAAGAG 2700
CAAGAGGAT CAACAAAAA GCGGATTTC TCTCAACGCT GGTAAACAGG CATATGGATG 2760
AGTTCAAGCG TACACAAAAG AATCTGGGAA ACTGGAAGA AGAAGCACAG CAGCTCTTAC 2820
AGAATGAAA AAGTGGGAGA GAGAAATCAG ATCAGCTGCT TTCCGTTGCC AATCTTGCTA 2880
AAGCAGAGC ACAGAGAGCA CTGAGTATGG GCAATGCCAC TTTTATGAA GTTGAGAGCA 2940
TCCTTAAAAA CCTCAGAGAG TTTGACCTGC AGGTGGACAA CAGAAAGCA GAAGCTGAAG 3000
AAGCCATGAA GAGACTCTCC TACATCAGCC AGAAGGTTT AGATGCCAGT GACAAGACCC 3060
AGCAAGCAGA AAGAGCCCTG GGGAGCGCTG CTGCTGATGC ACAGAGGCA AAGATGGGG 3120
CCGGGGAGGC CCTGGAAATC TCCAGTGAGA TTGAACAGGA GATTGGGAGT CTGAACCTTG 3180
AAGCCATAGT GACAGCAGAT GGAGCCTTGG CCATGGAAAA GGGACTGGCC TCTCTGAAGA 3240
GTGAGATGAG GGAAGTGGAA GGAGAGCTGG AAAGGAAGGA GCTGGAGTTT GACACGAATA 3300
TGGATGAGT ACAGATGGTG ATTACAGAAG CCCAGAAGGT TGATACCAGA GCCAAGAACG 3360
CTGGGGTTAC AATCTCAAGC ACCTCAACA CATTAGACGG CCTCTGCTCAT CTGATGGACC 3420
AGCCTCTCAG TGTAGATGAA GAGGGGCTGG TCTTACTGGA GCAGAAGCTT TCCCAGGCA 3480
AGACCCAGAT CACAGCCCAA CTGCGGCCCA TGATGTGAGA GCTGGAAGAG AGGGCAGGTC 3540
AGCAGAGGGG CCACTCCCTT TTGCTGGAGA CAAGCATAGA TGGGATTCTG GCTGATGTGA 3600
AGAACTTGA GACACTTAGG GACAACTGCG CCCCAGGCTG CTACAATACC CAGGCTCTTG 3660
AGCAACAGTG AAGCTGCCAT AAATATTCT CACTGAGGT TCTTGGGATA CAGATCTCAG 3720
GGCTCGGGAG CCAATGTCAT TGAGTGGGTG GGATGGGGAC ATTTGAACAT GTTTAATGGG 3780
TATGCTCAGG TCAACTGACC TGACCCCAT CTGTATCCCA TGGCCAGGTG GTTGTCTTAT 3840
TGCAACATAC TCCTTCTCTC CTGATGCTGG GCAATGAGGC AGATAGCACT GGGTGTGAGA 3900
ATGATCAAGG ATCTGAGCCC CAAAGAATAG ACTGGATGGA AAGACAACT GCACAGGCAG 3960
ATGTTTGCCT CATAATAGTC GTAAGTGAG TCTTGAAT TGGACAAGTG CTGTTGGGAT 4020
ATAGTCAACT TATTTCTTGA GTAATGTGAC TAAAGGAAAA AACTTTGACT TTGCCCAGGC 4080
ATGAAATCT TCTTAATGCT AGAACAGAGT GCAACCCAGT CACACTGTGG CCAGTAAAT 4140
ACTATTGCT CATATTGTCC TCTGCAAGCT TCTTGTCTGAT CAGAGTTCTT CCTACTTACA 4200
ACCCAGGGTG TGAACATGTT CTCCATTTC AAGCTGGAAG AAGTGAGCAG TGTGGAGTG 4260
AGGACCTGTA AGGCAGGCC ATTCAAGCT ATGGTGCTTG CTGGTGCTG CCACCTTCAA 4320
GTTCTGGACC TGGGCATGAC ATCTTTCTT TTAATGATGC CATGGCAACT TAGAGATTGC 4380
ATTTTATTA AAGCATTTCC TACCAGCAAA GCAATGTTG GGAAGTATT TACTTTTTCG 4440
GTTTCAAGT GATAGAAAAG TGTGGCTTGG GCATTGAAAG AGGTAAATTT CTCTAGATTT 4500
ATTAGTCTTA ATTCATCTCT TCTCTCTTTC CTCCACCCAT AATAAGAGAA TGTTCTTACT 4560
ATCTTATTTT CTCAATCTCC TCTCTCTTTC CTCCACCCAT AATAAGAGAA TGTTCTTACT 4620
CACACTTCAG CTGGGTCACA TCCATCCCTC CATTCATCTT TCCATCCATC TTTCCATCCA 4680
TTACCTCCAT CCTCTCTTCC AACATATATT TATTGAGTAG CTACTGTGTG CCAGGGGCTG 4740
GTGGGACAGT GGTGACATAG TCTCTGCCCT CATAGAGTTG ATTTGTAGT GAGGAAGACA 4800
AGCATTTTAA AAAAATAAAT TTAACCTTAC AAACCTTTGT TGTCAAGT GGTGTTTATT 4860
GCAATAACCG CTGGTTTGGC AACCTCTTTC CTCAACAGAA CATATGTTGC AAGACCTTCC 4920
CATGGGGCCA CTGAGTTTTC GGCAAGGCTG ACAGAGCTCT GGGTGTGCA CATTTCTTTG 4980
CATTCAGCT GTACATTTCT GCCTTTCTAC AACTGATTGC AACAGACTGT TGAGTTATGA 5040
TAACACCAAT GGAATTTGCT GGAGGAACCA GAGGCATTC CACCTTGGCT GGAAGACTA 5100
TGGTGCTGCC TTGCTTCTGT ATTTCTTGG ATTTCTCTGA AAGTGTTTT AAATAAGAA 5160
CAATTGTTAG ATGCC

Seq ID NO: 220 Protein sequence:
Protein Accession #: NP_005553

1 11 21 31 41 51
| | | | |

WO 02/086443

PCT/US02/12476

	MPALWLGCC	CPSLLLPAA	ATSREVCDC	NGKSRQCIFD	RELHRQTGNG	FRCLNCNDNT	60
	DGIHCECKN	GFYRHRERDR	CLPCNCSKG	SLSARCDNSG	RCSCCKPGVTG	ARCDRCCLPGF	120
	HMLTDAGCTQ	DQRLDLSKCD	CDPAGIAGPC	DAGRCVCCKPA	VTGERCDRCR	SGYINLDGNG	180
5	PEGCTQCPY	GHSASCRSSA	EYSVHKITST	FHQDVGWKA	VQRNGSPAKL	QWSQRHQDVP	240
	SSAQLDLPVY	FVAPAKPLGN	QQVSYGQSL	FDYRVDGRGR	HPSAHVDILE	GAGLRITAPL	300
	MPLGKTLPCQ	LTKYTTFRLN	EHPSNWSPQ	LSYFEYRRL	RNLTLALRIRA	TYGEYSTGYI	360
	DNVTLISARP	VSGAPAPWVE	QCICPVGYKG	QFCQDCASGY	KRDSARLGPF	GTCIPCNQCG	420
	GGACDPDTGD	CYSGDENPDI	ECADCPIGFY	NDPHDRPSCK	PCPCNHPFSC	SVMPEZEEV	480
10	CNCPGPGVTG	ARCELCAADG	FGDPFGEHGP	VRPCQPCQCN	NNVDPSASGN	CDRLTGRCLK	540
	CIHNTAGIYC	DQCKAGYFGD	PLAPNPADKC	RACNCPMGS	EPVGCSDGT	CVCKPGFPGP	600
	NCEHGAFCSC	ACYNQVKIQM	DQFMQQLQRM	EALISKAQGG	DGVVPDTELE	GRMQAEQAL	660
	QDILRDAQIS	EGASRSGLQ	LAKVRSQENS	YQSRLLDLKM	TVERVRALGS	QYQNRVRDTH	720
	RLITQMQLSL	AESEASLGNT	NIPASDHYVG	PNGFKSLAQE	ATRLAESHVE	SASNMEQLTR	780
15	ETEDYSKQL	SLVRKALHEG	VSGSGSPDQ	AVVQGLVEKL	EKTKSLAQQL	TREATQAEIE	840
	ADRSYQHSLE	LLDSVSRLOQ	VSDQSFQVEE	AKRIKQKADS	LSTLVTRMD	EPKRTQKHLQ	900
	NWKEEAQQL	QNGKSGREKS	DQLLSRANLA	KSRQAELSM	GNATFYEVES	ILKNLREFDL	960
	QVDNRKAEAE	EAMKRLSYIS	QKVSADSDKT	QQAERALGSA	AADAQRAKNG	AGEALEISSE	1020
	IEQEIGSLNL	EANVTADQAL	AMEKGLASLK	SEMREVEGEL	ERKELEPDIN	MDAVQMVITE	1080
20	AQKVDTRAKN	AGVTIQDTLN	TLDGLHLMD	QPLSVDEEGL	VLEQLKLSRA	KTQINSQLRP	1140
	MMSLEERAR	QQRGHLHLE	TSIDGILADV	KNLENIRNKL	PPGCYNTQAL	EQQ	

Seq ID NO: 221 DNA sequence
Nucleic Acid Accession #: NM_016529
Coding sequence: 13-1854

25	1	11	21	31	41	51	
	GTCAAGAAAA	GAATGTCGT	AATTGTTTGA	ACTCCTTCAG	GACGACTTCG	GCTTTACTGT	60
	AAAGGGGCTG	ATAATGTGAT	TTTTGAGAGA	CTTTCAAAAG	ACTCAAAATA	TATGGAGGAA	120
30	ACATTATGCC	ATCTGGGAATA	CTTTGCCACG	GAAGGCTTGC	GGACTCTCTG	TGTGGCTTAT	180
	GCTGATCTCT	CTGAGAATGA	GTATGAGSAG	TGGCTGAAAG	TCTATCAGGA	AGCCAGCACC	240
	ATATTGAAG	ACAGAGCTCA	ACGGTTGGAA	GAGTGTTCAG	AGATCATTGA	GAAGAATTGT	300
	CTGCTACTTG	GAGCCACAGC	CATAGAAGAT	CGCCTTCAAG	CAGGAGTTCC	AGAAACCATC	360
	GCAACACTGT	TGAAGTCGCA	AATTAATAA	TGGGTGTTGA	CAGGAGACAA	ACAAGAAACT	420
35	GCGATTAATA	TAGGGTATTC	CTGCCGATTG	GTATGCGAGA	ATATGGCCCT	TATCCTATTG	480
	AAGGAGGACT	CTTTGGATGC	CACAAGGGCA	GCCATTACTC	AGCACTGCAC	TGACCTTGGG	540
	AATTGCTCTG	GCAAGGAAAA	TGACGTGGCC	CTCATCATCG	ATGGCCACAC	CCTGAAGTAC	600
	GCGCTCTCCT	TGCAAGTCOG	GAGGAGTTTC	CTGGATTGGG	CACTCTCTGT	CAAAGCGGTC	660
40	ATATGCTGCA	CGAGTGTCTC	TCTGCAGAAG	TCTGAGATAG	TGGATGTGGT	GAAGAAGCGG	720
	GTGAAGGCCA	TCACCTCTGC	CATCGGAGAC	GGCGCCACAG	ATGTGGGGAT	GATCCAGACA	780
	GCCCACTGTG	GTGTGGGAAT	CAGTGGGAAT	GAAGGCATGC	AGGCCACCAA	CAACTCGGAT	840
	TACGCCATCG	CACAGATTTT	CTACTTAGAG	AAGCTTCTGT	TGTTTCATGG	AGCCTGGAGC	900
	TACAACCGGG	TGACCAAGTG	CATCTTGATC	TGCTTCTATA	AGAACGTGGT	CCTGTATATT	960
45	ATTGAGCTTT	GGTTCGCCCT	TGTTAATGGA	TTTTCTGGGC	AGATTTTATT	TGAAGCTTGG	1020
	TGCATCGGCC	TGTACAATGT	GATTTTCACC	GCTTTGCCGC	CCTTCACTCT	GGGAATCTTT	1080
	GAGAGGTCTT	GCATCAAGGA	GAGCATGCTC	AGSTTTCCCC	AGCTCTACAA	AATCAACCCG	1140
	AATGGCGGCA	GCTTCAACAC	AAAGGTTTTT	TGGGGTCACT	GCATCAACGC	CTTGGTCCAC	1200
	TCCCTCATCC	TCTTCTGGTT	TCCCATGAAA	GCTCTGGAGC	ATGATACTGT	GTTTGACAGT	1260
	GGTCACTGCA	CGAGTATTAT	ATTGTTGGGA	AATATTGTTT	ACACATATGT	TGTTGTGACT	1320
50	GTTTGTCTGA	AAGCTGGTTT	GGAGACCACA	GCTTGGACTA	AATTCAGTCA	TCTGGCTGTC	1380
	TGGGGAAGCA	TGCTGACCTG	GCTGGTGTTT	TTTGGCATCT	ACTCGACCAT	CTGGCCCAAC	1440
	ATTCCCATGG	TCCAGATAT	GAGAGGACAG	GCAACTATGG	TCTCGAGCTC	CGCACACTTC	1500
	TGGTTGGGAT	TATTTCTGGT	TCTTACTGCC	TGTTTGATTG	AAGATGTGGC	ATGGAGAGCA	1560
55	GCCAAAGACA	GCTGCAAAA	GACATTGCTG	GAGGAGGTGC	AGGAGCTGGA	AACCAAGTCT	1620
	CGAGTCTCTG	GAAGAGCGGT	GCTGCGGGAT	AGCAATGGAA	AGAGGCTGAA	CGAGCGCGAC	1680
	CGCCTGATCA	AGAGGCTGGG	CGGGAAGACG	CCCCGACGCG	TGTTCCGGGG	CAGCTCCCTG	1740
	CAGCAGGGCG	TCCCGCATGG	GTATGCTTTT	TCTCAAGAAG	AACACGGAGC	TGTTAGTCAG	1800
	GAAGAAGTCA	TCCGTGCTTA	TGACACCACC	AAAAAGAAAT	CCAGGAAGAA	ATAAGACATG	1860
60	AATTTTCTTG	ACTGATCTTA	GGAAAGAGAT	TCAGTTTGT	GCACCCAGTG	TTAACACATC	1920
	TTTGTGAGAG	AAGACTGGCG	TCCAAGGCCA	AAACACCAGG	AAACACATTT	CTGTGGCCTT	1980
	AGTTAAGCAG	TTTGTGATTT	ACATATTTCC	TGCGAAACCT	GGAGTGCAGA	CCACAGGGGA	2040
	AGCTATCTTT	GCCCTCCCAA	CTCGTCTGCA	GTGCTTAGCC	TAACTTTTGT	TTATGTCCGT	2100
	ATGAAGCATT	CAACTGTGCT	CTGTGAGGTC	TCAAATTAAA	AACATTATGT	TTCACCAATA	2160
65	AGAAAAA	AAAAA					

Seq ID NO: 222 Protein sequence:
Protein Accession #: NP_057613

70	1	11	21	31	41	51	
	MSVIVRTPSG	RLRLYCKGAD	NVIPERLSKD	SKYMEETLCH	LEYFATEGLR	TLCVAYADLS	60
	ENEYEENLKV	QGEASTILKD	RAQRLEECYE	IEKNLLLLG	ATAIEDRLQA	GVPETIATLL	120
	KAEIKIWLVT	GDKQETAINI	GYSCLRVSON	MALILLKEDS	LDATRAAITO	HCTDLGNLLG	180
75	KENDVALIID	GHTLKYALSF	EVRRSFLDLA	LSCKAVICCR	VSPLOKSEIV	DVVKKRVKAI	240
	TLAIGDGAND	VGMQTAHVQ	VGISGNEGMQ	ATNNSDYAIA	QPSYLEKLLL	VHGAWSYNRV	300
	TKCILYCFYK	NVLYIIEIWL	FAFVNGFSQO	ILFERWICGL	YNVIFTALPP	FTLGIFERSHC	360
	TOESMLRFPQ	LKTIQNGEG	FNTKVPWGHG	INALVHSLIL	FWFPMKALEH	DTVFDSEHAT	420
	DYLFVGNIVY	TYVVVTCLRK	AGLETTAWTK	FSHLAVWGSN	LTLVFPFIY	STIWPITPIA	480
80	PDMRGQATMV	LSSAHFWLGL	FLVPTACLIE	DVAWRAAKHT	CKKTLLEEVO	ELETKSRVLG	540
	KAVLRDSNGK	RLNERDRLIK	RLGRKTPPTL	FRGSSLLQGV	PHGYAFSQEE	HGAVSQEEVI	600
	RAYDTTKKKS	RKK					

Seq ID NO: 223 DNA sequence
Nucleic Acid Accession #: BC017001
Coding sequence: 1-394

85	1	11	21	31	41	51	
----	---	----	----	----	----	----	--

5
10
15
20
25
30
35
40
45

```

AACGCTGGGC AGGGCCGGCG CGGGTCGGGG GGGCCCGAG GGGCCCGGC CGAGCGGGG 60
CGGCGAGGGC GGCAGCATCC ACTCGGGCCG CATCGCCGCG GTGCACAACG TGCCGCTGAG 120
CGTGCTCATC CGGCCCGTCG CGTCCGTGTT GGACCCCGCC AAGGTGCAGA GCCTCGTGA 180
CACGATCCGG GAGGACCCAG ACAGCGTGCC CCCCATCGAT GTCCCTCTGA TCAAAGGGGC 240
CCAGGAGGAT GACTACTTCT ACTCCTTTGG GGGCTGCCAC CGCTACGCGG CCTACCAGCA 300
ACTGCAGCGA GAGACCATCC CGGCCAAGCT TGTCCAGTCC ACTCTCTCAG ACCTAAGGGT 360
GTACCTGGGA GCATCCACAC CAGACTTGCA GTAGCAGCCT CCTTGGCACC TGCTGCCACC 420
TTCAAGAGCC CAGAAGACAC ACCTGGCCTC CAGCAGGCTG GGCATGCAG AAGGGATAGC 480
AGGGGTGCAT TCTCTTTGCA CCTGGCGAGA GGGTCTGACT CTGGGCACCC CTCTCACCAG 540
CTACAAGGCC TTGGACTCAC TGTACAGTGT GGGAGCCCCA GTTCCCACCT CTGTGACAA 600
AGGATCATGG CCTTACCCCT GAAGCATTAC CGAGAAGGAG AACAGAGATG GCCTTGAAGA 660
GCCACGTGCT GCGCGCTCCA AATTCCCAAG GACAAGGATC CCTCTGCATT TTTGTCTATG 720
TAACCTCTTA TATGGACTAC ATTCAGCTGC AAGGAAGGA AAACCTTGAT TGCAGTGGTT 780
TAACAAACAA GAAGATTGTT TTCCACATA GCATGGATTG TGGAGATGGG TGGCTAATGG 840
TATTGGTTCA ACACCTCCAC GGAGGTAGGG GTACAGCTCT GGATCCTTTT GCCTTAATCT 900
CAGTGCTCGT TACTTCATGG TCCCAAGATG GCTGCTGTAT CCCCAGAAT CATGCTCGCG 960
TTCAAGGAAG GAGGGTGGGA GGAAGAGGAA GGGCCAAACT AGCTGGACCC GTCCCTTCT 1020
ATCAGAAAGT AAAACCTCGT CAGAAGTCTG TTTCTGCTC TCTCCCTCTG CATATCTTCA 1080
CTTAGATGCC CTGTGGCCGA GCCAGCTACC ATTGCACCTC TAGCTGCAAA CAAAGCTAAG 1140
ACAGCAGGGA ACAGAATTGT CATGGCTGAA TAGACCAATC GTGTTCCATC TACTGAGACT 1200
GGCACACTGT TACTCTGCAAT AAAACTGGGA TCCCATTAAC AAGAGAGAAA TGCAGAATTG 1260
TGTACCAAGT AGCTTTTGCT GTGTAACAAA CCATCCCCAA ACTTGGCAGC TAGAAACAAA 1320
CCCTGTATT TCCCACAATC CTATGGGTTG GCAATTTGGG CTGGGCTCAA CAGGGCAGTT 1380
CTGCTGCTCA CACCTGGGAT CCCTCATGGA GCTAAGGTCA GCTGTACCT CAGCTGGGCC 1440
TGGATGGTCT AGGATAGCCT TACTCACTTG CCTGGCAGGT GACAGGCTGT TGGCTGGAAT 1500
TGCTTGGTTC TCTCTCATGT GGCCTCTCCA GCAGGCTAGC TCAGGCTTAT TCACATGATG 1560
GCTTCAGGAT TCCAAAGAGA GTGAGAGTAG AAGCTGAAAG ACTTCTTGAG TTCTTGGCCT 1620
GGAAGTGGGA CTAGGACAGT GTCACTTCTG CTAAGTTCTT TTGGTCAGAG CAAATCACA 1680
GGCTTTACCC AGATTCAAGG GATGAGAAAC AGACTACATG TCTTGATGAG GGGAAACCACA 1740
AAGAGCTTGT GGCCATTTTT CACCTATCAC AAATAATTTT GGATGGGTAT TTATTGGAT 1800
AAAGGTATTT CCTTCTTCCC CTGGCCCTCA GGGTCTAGGG GAAGAGGTGT TGGGGCAGGA 1920
GAAGGCACTA AGACATTGTC CTGGCCCTCA GGGTCTAGGG GAAGAGGTGT TGGGGCAGGA 1980
AGTGAGTCTC TCCATGGGCT GGAACCACTG TAGTAGGAGT GCCTCCTTGT CTGCACTGCT 2040
GGTATGGGCT GTATGGGCT AGGACATTCC AGAGGGGCTT CTGAAAACCA AGAGTCCCTG 2100
GGGAAGGGA ACAGAGTAAG GCAGGCTTGC TTCTCACTGC CCTCTAAGGG AACTTGGTCA 2160
CTCGGCACCT TTAAGCCTCA GTTCTCCAG TTCAATAATA AGGACAAGAG CTTTCCCAT 2220
GCATTCTCTT TCCCGGGGAA AGTTGACTGA GGTGACCAAT AATAGAATTG AAAAGGGAGA 2280
GTGCTCTCAG TGCAATGTGG CATCTGGAT TGGGTCTTGG AACAAAACA GGACATTAGT 2340
GGGAAATG GAAATCTGAA AAAAGTCTGA ATTTAGTTA ATATACCAAT TTCAGTCTCT 2400
TGGTTTGAC AGATGTACCA TGGTGTGTA AGATGTGAC CTGGGGTAG GCTGGGTGAA 2460
GGGTATACAG GAACTCTTTG TACTATCTCT GCAACTTCTC TGTAATCTA GTATCATTCC
AAAAATAAAG TTTATTTAAT TAAAAAAA AAAAATAA AA

```

Seq ID NO: 224 Protein sequence:
Protein Accession #: AAH17001.1

50

```

1 11 21 31 41 51
TLGRAGAGRG APEGPGPSGG AQGGSIHSGR IAAVHNVPLS VLIRPLPSVL DPAKVQSLVD 60
TIREDDPSVP PIDVLWIKGA QGGDYFYSPG GCHRYAAQQ LQRETIPAKL VQSTLSDLRV 120
YLGASTPDQ

```

55
Seq ID NO: 225 DNA sequence
Nucleic Acid Accession #: NM_021048
Coding sequence: 1..1110

60
65
70
75
80

```

1 11 21 31 41 51
ATGCTCTGAG CTCCAAAGCG TCAGCGCTGC ATGCTCTGAAG AAGATCTTCA ATCCCAAAGT 60
GAGACACAGG GCCTCGAGGG TGCACAGGCT CCCCTGGCTG TGGAGGAGGA TGCTTCATCA 120
TCCACTTCCA CCAGCTCCTC TTTTCCATCC TCTTTTCCCT CCTCTCTCTC TTCTCTCTCC 180
TCCCTCTCTC ATCCTCTAAT ACCAAGCACC CCAGAGGAGG TTCTGCTGTA TGATGAGACA 240
CCAAATCCTC CCCAGAGTGC TCAGATAGCC TGCTCTCTCC CCTCGTCTGT TGCTTCCCTT 300
CCATTAGATC AATCTGATGA GGGCTCCAGC AGCCAAAAGG AGGAGAGTCC AAGCACCCCTA 360
CAGGTCTGTC CAGACAGTGA GTCTTTACCC AGAAGTGAGA TAGATGAAAA GGTGACTGAT 420
TTGGTGCAGT TTCTGCTCTT CAAGTATCAA ATGAAGGAGC CGATCACAAA GGCAGAAATA 480
CTGGAGAGTG TCATAAAAAA TTATGAAGAC CACTTCCCTT TGTGTTTAG TGAAGCCTCC 540
GAGTGCATGC TGCTGGTCTT TGGCATTGAT GTAAAGGAAG TGGATCCAC TGGCCACTCC 600
TTTGTCTCTG TCACCTCCTT GGGCCTCACC TATGATGGGA TGCTGAGTGA TGTCCAGAGC 660
ATGCCCAAGA CTGGCAATCT CATACCTATC CTAAGCATAA TCTTCATAGA GGGCTACTGC 720
ACCCCTGAGG AGGTCTATCT GGAAGCACTG AATATGATGG GGCTGTATGA TGGATGGAG 780
CACCTCATTT ATGGGGAGCC CAGGAAGCTG CTCACCCAAG ATTGGGTGCA GGAAGACTAC 840
CTGGAGTACC GGCAGGTGCC TGGCAGTGAT CCTGCACGGT ATGAGTTTCT GTGGGTCCA 900
AGGGCTCATG CTGAAATTAG GAAGATGAGT CTCCTGAAAT TTTTGGCCAA GGTAAATGGG 960
AGTGATCCAA AGTCTCTCCC ACTGTGGTAT GAGGAGGCTT TGAAGATGA GGAAGAGAGA 1020
GCCACGAGCA GAATGCGCAC CACAGATGAT ACTACTGCCA TGGCCAGTGC AAGTCTAGC 1080
GCTACAGGTA GCTTCTCTCA CCTGAATAA

```

Seq ID NO: 226 Protein sequence:
Protein Accession #: NP_066386

85

```

1 11 21 31 41 51
MPRAKPRQR MPEDLQSQS ETQGLEGAQA PLAVREDASS STSTSSSFPS SPPSSSSSSS 60

```

SSCYPLIPST PEEVSADDET PNPPQSAQIA CSSPSVVASL PLDQSDGESS SQKEESPSTL 120
QVLDPDESLEP RSEIDEKVID LVQFLLFKYQ MKEPIITKAEI LESVIKNYED HFPLLFSEAS 180
ECMLLVFGID VREVDPTGHS FVLVTSGLT YDGMLSDVQS MPKTGILILI LSIIFIEGYC 240
TPEEVIWEAL NMMLGYDME HLIYGEPRKL LTQDWVQENY LEYRQVPGSD PARYEFLWGP 300
RAHAEIRKMS LLKFLAKVNG SDPRSFLWY EEALKDEER AQDRIATDD TTAMASASSS 360
ATGSFSYPE

Seq ID NO: 227 DNA sequence
Nucleic Acid Accession #: NM_005025.1
Coding sequence: 82-1314

1 11 21 31 41 51
GCGGAGCACA GTCCGCGCAG CACAAGCTCC AGCATCCCGT CAGGGGTGTC AGGTGTGTGG 60
GAGGCTTGAA ACTGTTACAA TATGGCTTTC CTGGACTCTT TCTCTTGTCT GGTTCCTGCA 120
AGTATGGCTA CAGGGGCCAC TTCCCTGAG GAAGCCATTG CTGACTTGTC AGTGAATATG 180
TATAATCGTC TTAGAGCCAC TGGTGAAGAT GAAATATTC TCTCTCTCC ATTGAGTATT 240
GCTCTTGCAA TGGGAATGAT GGAACCTGGG GCCCAAGGAT CTACCCAGAA AGAAATCCGC 300
CACTCAATGG GATATGACAG CCTAAAAAT GGTGAAGAAT TTTCTTCTT GAAGGAGITT 360
TCAAACATGG TAACCTGCTAA AGAGAGCCAA TATGTGATGA AAATGCCAA TTCCTTGTIT 420
GTGC AAAATG GATTTCATGT CAATGAGGAG TTTTGTCAAA TGATGAAAAA ATATTTTAA 480
GCAGCAGTAA ATCATGTGGA CTTCAGTCAA AATGTAGCCG TGGCCACTA CATCAATAAG 540
TGGGTGGAGA ATACACAAA CAATCTGGTG AAAGATTGG TATCCCCAAG GGATTTTGAT 600
GCTGCCACTT ATCTGGCCCT CATTAAATGCT GTCTATTTC AAGGGAACTG GAAGTCGAG 660
TTTAGCCCTG AAAATACTAG AACCTTTTCT TTCCTAAAG ATGATGAAAG TGAAGTCCAA 720
ATTCGAATGA TGATACAGCA AGGAGAATTT TATTATGGGG AATTGTAGTGA TGGCTCCAA 780
GAAGCTGGTG GTATCTACCA AGTCCTAGAA ATACCATATG AAGGAGATGA AATAAGCATG 840
ATGCTGGTGC TGTCCAGACA GGAAGTTCTT CTGTCTACTC TGGAGCCATT AGTCAAGCA 900
CAGCTGGTGC AAGAATGGGC AAATCTGTGT AAGAAGCAAA AAGTAGAAGT ATACCTGCCC 960
AGGTTACAG TGGAACAGGA AATTGATTTA AAAGATGTTT TGAAGGCTCT TGAATAACT 1020
GAAATTTTCA TCAAAGATGC AAATTTGACA GGCCTCTCTG ATAATAAGGA GATTTTCTT 1080
TCCAAAGCAA TTCAAGAGT CTCTCTAGAG GTTAATGAAG AAGGCTCAGA AGCTGCTGCT 1140
GTCTCAGGAA TGATTGCAAT TAGTAGGATG GCTGTGCTGT ATCCTCAAGT TATTGTGAC 1200
CATCCATTTT TCTTTCTTAT CAGAAACAGG AGAAGTGGTA CAATTCATT CATGGGACGA 1260
GTCTGCTATC CTGAACAACT GAACACAACT GCACATGATT TCGAAGAACT TTAAGTTACT 1320
TTATTTGAAT AACAAAGAAA ACAGTAACCT AGCAGATTAT GTTTGCAACT GGTATATATT 1380
TAGGATTGTT GTTTTACAGT ATATCTTAAG ATAATATTTA AAATAGTTCC AGATAAAAC 1440
AATATATGTA AATATATAGT AACTTGTCAA GGAATGTTAT CAGTATTAG CTAATGGTCC 1500
TGTTATGTCA TTGRTTGTGT GTGCTGTGT TAAAAATAAA AGTACCTATT GAACATGTG

Seq ID NO: 228 Protein sequence:
Protein Accession #: NP_005016.1

1 11 21 31 41 51
MAFLGLFSL VQSMATGAT PPEEAIADLS VNMYNRLRAT GEDENILFSP LSIALAMGMM 60
ELGAQGSTQK EIRHSMGYDS LKNGEFSFL KEFSNMVTAK ESQVVMKIAN SLFVQNGFHV 120
NEEFLOMKK YPNAAVNHVD FSNVAVANY INKVENNTN NLVKDLVSPR DFDAATYLAL 180
INAVYFKGNW KSQPRPENTR TFSFKDDDES EVQIPMMYQQ GEFYVGEFSD GSNAGGIYQ 240
VLEIPVEGDE ISMMLVLSRQ EVPLATLEPL VKAQLVEEWA NSVKKQKVEV YLPRFTVEQE 300
IDLKDLKAL GITEIFIKDA NLTLGSDNKE IFLSKAIHKS FLEVNEEGSE AAASVGMIAI 360
SRMAVLPOV IVDEPPFFLI RNRRTGTILF MGRVMHPETM NTSGHDFEEL

Seq ID NO: 229 DNA sequence
Nucleic Acid Accession #: NM_003695
Coding sequence: 12-398

1 11 21 31 41 51
CGACATCAGA GATGAGGACA GCATTGCTGC TCCTTGACGC CCTGGCTGTG GCTACAGGGC 60
CAGCCCTTAC CCTGCGCTGC CACGTGTGCA CCAGCTCCAG CAACTGCAAG CATTCTGTGG 120
TCTGCCCAGG CAGCTCTCGG TTCTGCAAGA CCACGAACAC AGTGGAGCCT CTGAGGGGGA 180
ATCTGGTGAA GAAGGACTGT GCGGAGTGGT GCACACCCAG CTACACCCCTG CAAGGCCAGG 240
TCAGCAGCGG CACCAGCTCC ACCCAGTGCT GCCAGGAGGA CCTGTGCAAT GAGAAGCTGC 300
ACACAGCTGC ACCCACCOCG ACCGCCCTCG CCCACAGTGC CCTCAGCCTG GGGCTGGCCC 360
TGAGCCTCCT GGCCTCATC TTAGCCCCCA GCCTGTGACC TTCCCCCAG GGAAGGCCCC 420
TCATGCTTTT CCTTCCCTTT CTCTGGGGAT TCCACACCTC TCTTCCCCAG CCGGCAACGG 480
GGGTGCCAGG AGCCCCAGGC TGAGGGCTTC CCCGAAAGTC TGGGACCAGG TCCAGGTGGG 540
CATGGAATGC TGATGACTTG GAGCAGGCC CACAGACCCC ACAGAGGATG AAGCCACCCC 600
ACAGAGGATG CAGCCCCCAG CTGCATGGAA GGTGGAGGAC AGAAGCCCTG TGGATCCCCG 660
GATTTCACAC TCCTTCTGTT TTGTTGCGGT TTATTTTGTG CTCAAATCTC TACATGGAGA 720
TAAATGATT AAAC

Seq ID NO: 230 Protein sequence:
Protein Accession #: NP_003686

1 11 21 31 41 51
MRTALLLLAA LAVATGPALT LRCHVCTSSS NCKHSVVCPA SSRPCKTNT VEPLRGNLVK 60
KDCAESCTPS YTLQGVSSG TSSTQCQED LCKNEKLHNA PTRTALAHS LSLGLALSLL 120
AVILAPSL

Seq ID NO: 231 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 126-752

	1	11	21	31	41	51	
	CCGGGCGAGGT	GGCTCATGCT	CGGGAGCGTG	GTTGAGCGGC	TGGCGCGGTT	GTCCTGGAGC	60
	AGGGGCGCAG	GAATTTCTGAT	GTGAACTAA	CAGTCTGTGA	GCCCTGGAAC	CTCCACTCAG	120
5	AGAAGATGAA	GGATATCGAC	ATAGGAAAAG	AGTATATCAT	CCCCAGTCCT	GGGTATAGAA	180
	GTGTGAGGGA	GAGAACCCAGC	ACTTCTGGGA	CGCACAGAGA	CGTGAAGAT	TCCAAGTTCA	240
	GGAGAACTCG	ACCGTTGGAA	TGCCAAGATG	CCTTGGAAAC	AGCAGCCCGA	GCCGAGGGCC	300
	TCTCTCTTGA	TGCTCCCATG	CATTCTCAGC	TCAGAATCCT	GGATGAGGAG	CATCCCAAGG	360
10	GAAAGTACCA	TCATGGCTTG	AGTGCTCTGA	AGCCCATCOG	GACTACTTCC	AAACACCAGC	420
	ACCCAGTGGG	CAATGCTGGG	CTTTTCTCT	GTATGACTTT	TTCTGCGCTT	TCTTCTCTGG	480
	CCCGTGTGGC	CCACAAGAAG	GGGAGCTCT	CAATGGAAGA	CGTGTGCTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAGAA	GAGCTGAATG	600
	AAGTTGGGCC	AGACGCTGCT	TCCCTGCGAA	GGGTTGTGTG	GATCTTCTGC	CGCACCAGGC	660
15	TCATCTGATC	CATCGTGTGC	CTGATGATCA	CGCAGCTGSC	TGGCTTCAGT	GGACCAAAAT	720
	TTCAAGATGG	CTGTATTCTG	CGGTGAGAAT	GAGAGAGTCA	AGCTGGGCAG	AATCTCTCGC	780
	CAAGAGTTCA	GCTTCTCTTT	GGAGACTGCT	CCATCAGTGC	CGAGGTGTGT	GGGAACAGGC	840
	TTCACTGCAC	CGCCATCTTA	CTGAGTTGCT	TCACGTGAGG	AAAAGGGGGC	TTTGGCCCTG	900
	TGACTCAGTT	CCACATTTTG	GATTGCATAC	TGGAAAAGAA	GCCATCTTTC	TTGCTAGTAA	960
20	ACCAGCAACC	CGGCTGTATA	CAGTGGTGAC	CCAAGCAATG	GATATAAAC	TAAAAATCTG	1020
	AGGAGGGGGA	GAGGTGGAAT	ACAGTAGTTC	TTGGAATCTG	AAGTCTCTTA	TTGATCAGG	1080
	TTATTTCTTG	GGACTTGGCA	AAAACTGAT	TGGTGGGGAT	CTCCTAGGAC	CTAGTGGACA	1140
	TCTGGTATTA	ATTTAATCTC	AGGAAAAACA	AGAAATTAAC	CCAGAGAGAG	TCGGGTTTTT	1200
	GGAAATCAGC	GTACCTACCT	CCAGACCGTG	GTGCTGTGCC	TCCATTTTTC	TCTGTCTATC	1260
25	AGCTCTGACT	TACGCTGACA	GTCACTTTG	CTATAAGGCA	CCTGGGTAGA	AGGGTGGATG	1320
	GGCTTCACAT	CAATTTTTTT	CTTCTTTAG	GGTGGGGGAT	TGGTTTGCT	TTCTTTTGTT	1380
	GTGGTTTTTT	TTTTTATTTT	TGTCAAGATT	GATTTTTAGA	TGCAAGGACT	TGAAAAGACC	1440
	CAGAAGGATG	CACCAAGTTT	TTCTTTGAGG	CCTAGGATTT	TTTATCTGCT	CCGAGCAGA	1500
	GGTAATTCCT	CACAACTTAT	TGCACCAATA	GCACAGCCCA	TTTTGAGCAG	AGTACCTCTT	1560
30	TGGGAGGCTT	TTGCTTTTGT	TTTGTTTTGA	ATTCTCTTTC	CTTAGCAGCA	AGGTCTTTTT	1620
	TCCTAGAGAA	TCTACTCGCT	TGCAGATCA	TTGCAACCTC	AGGAGCCCTC	ACTGATTGAG	1680
	TGCTGTGAGC	CTGATATACT	ACTTTGGACT	CTGGAACACG	ATATGGGCTC	TATTTCTCTAT	1740
	TTCTACTGTG	TGTGTTTAA	CAACCGTCGG	AGACCAGATG	ACCTGTGAGA	TGGCTAGTCC	1800
	TGTATAACTC	GACTCTGTAT	GTTTCAATGT	ATGTTACTGC	AATGCTTCAC	CTGCTGTACA	1860
35	GTGTTTGTGA	GATGCTCTTT	GAAGATGGTA	CTTTTATATT	T		

Seq ID NO: 232 Protein sequence:
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
	MKDIDIGKEY	IIPSPGYRSV	RERTSTSGTH	RDREDSKFR	TRPLEQDQAL	ETAARAEGLS	60
	LDASMHSQLR	ILDEEHPKAG	YHGLSALKP	IRTTSKHQHP	VDNAGLFSCM	TPSWLSLAR	120
	VAHKKGELSM	EDVWSLSKHE	SSDVNCRRL	RLWQEELNEV	GPDAASLRRV	VWIFPCTRLL	180
45	LSIVCLMITQ	LAPSPGNPNQ	DGCILRSE				

Seq ID NO: 233 DNA sequence
Nucleic Acid Accession #: CAT cluster

	1	11	21	31	41	51	
	TTTAAATGGT	GCTCATATAT	ACTGTATTTT	TTGTTGTTTA	GTTTTACTTA	TTGAGAGTGT	60
	CACACATGTA	ATCACAATAA	CATGATTTTT	TTTTTTTACT	TTTACTCCCC	AAATTAATCA	120
	TGTTTCTTAG	ATCGTAGTCA	TTGAGAAGTC	CCAATAACTC	TAAACTTTTG	AGTTATAACG	180
	TAGTAAACTT	CTCTTTCAT	TTGTGTTAG	CTCTGTAGTC	TAACTCTGGA	TTTAAATTTT	240
55	TTTGTTTCCA	AAGTCACAAT	TGAATTATTC	TTAGATACCT	TAAGCCACTG	AATTCAGTTC	300
	TGTTTGTACT	AAAGCAAAAC	AACGTGACAG	TTTATTTTCA	AACTACTACT	TCTTGATATT	360
	TTGTTATGGT	ATATCTTTTT	ATTAAATATT	TATTTTGACT	AAGCTTTTCA	AAAATATTTG	420
	AAGCTATTTT	AATCATCAAG	TATGAAAAC	AAATTAATCT	TGCATTTTCC	TATATATGCA	480
60	TATATTATGG	ATTAAACAGA	ATTGTATCAT	TTTTGGCCTA	ATGCTGAGAT	ATAAAAGATA	540
	ATTAGCCTAC	TATAGTATTA	ATAAATTTTT	CAGTTGGTTT	GGGCAAAATT	AAACCTGAAA	600
	AATAGGTTAA	AAAGTAGTTA	CAAAATAAAC	TTACTAATTT	ATACCTGATT	TTTTTCTTTG	660
	AATTAAAGTA	CATTTTAAAT	GAGCTTTATA	ATACCTTAAA	AAGTTGGTTC	TAAATTAATA	720
	TATGAAAGCT	CTGGCTATCA	TCTGGGATA	GTAATTTCTA	ATTATATAGT	ATTCAAAAC	780
	TATATATTTT	TTAGTTCCTT	TGAGATAACT	AATTTCTAAT	TATATATGTT	TCAAAAACCA	840
65	TATCCTGTAT	TTTTTTTAA	AATTGTTTTA	TAAATAGGTC	ATAAGATACA	AGGTCTGCAT	900
	TAGAAGACCC	ACTCTTACTA	GGTCCCTTAA	GGATCTGCCA	TAGATTTTTT	TTTTTTTTTT	960
	TTTTTTTTAG	GTAGTTTAAA	GCAAGCACTG	ATACCACTGG	GAGTTGGTCT	TGATCTAGGA	1020
	GATTCGTGTA	AGCATCCAAA	AACAATGCCT	AATTTCAAGT	CTTAGGTTAT	GGCTGTGAC	1080
70	TCCAGATAAA	AGATGGAGAA	TACCTCATGT	ACTGTGACTT	GAAAATGAAT	TCTTAAATTT	1140
	CTTAGGCTCT	CTCCATGTAT	CTTCTTAAAG	GAAAAGTTTC	TGAGTGTGAT	CTCTCTTTTG	1200
	CCATAGTATC	AAGTGGAGGG	TAGTTCAGAA	AAGTTAATAG	GAAATCTTTT	GTGACAGCAG	1260
	ACTATAATAG	AAGTTTGAGT	AATATTTTAA	TAAATTTATA	TAATTCAAAT	GATAAAATG	1320
	TATCAATGTT	ATCCAATGAT	TTTTATTAAA	AAATTACCTT	ATTATTAGAA	CTGTGCCTAT	1380
	TACATAAAAA	GTGCTCATGT	ATTGGAATTT	TAAATAATTT	ATTAAATCA	AGACCACCAT	1440
75	AAGTCATTAA	TAAATTAATA	ATTGTTTTAA	ATCAGTGGTT	TTCAACCTTC	ACTTCATATT	1500
	AGAATCATCT	GAGGACTTTT	AATATGGAAT	CCACCTCAT	ACAATTAAGT	CTAAATTTCT	1560
	GGAAGATGGA	GCCATGCTTG	TTTTTCCAAA	AGCTCTTTGA	GTGATTCTAA	TTGTAGTCA	1620
	GAGTTGAAGA	CCACTGCTCT	AAATTAGTGC	AGGAAAATGC	TTTTATTCTT	CCCATGTTAA	1680
	CTTTTAAAC	TAGTAATGTA	CCCAGTTAAG	TTTTGATGGT	TTAAATTTCA	CTAAGAAACA	1740
80	TATTCCTCTA	ATAACTAGCA	TTTATTACAT	GAAATTTAAG	AGTTTAAAGT	CCATCAAACT	1800
	AGCCCTTGTG	TAGATATTAT	ATTCTTCTC	TATAACTTCA	AAATAGATAT	TTCAATCAAA	1860
	CTGTCTCAGT	GAGAAACAT	AATGGAATTT	TTTTTTTTTC	CTCTGGAGCT	GCTGTTCAG	1920
	TGAGATGGAG	GAGGTGGGCA	CATTTAAGGT	CAGTTCACCT	ACCTATGGTT	CAGAGTTCTG	1980
	ATCATATTGA	AGTTTGGAAA	AGAGAGCTTA	TCACAGGTTT	GTATGCTGGT	GAATGGATAG	2040
85	TTTAAATTTT	CACGTCTCTA	AAAGAGAACT	AGCTCTCCAG	CAGTCTCAGA	AAAGCTTTGA	2100
	CAATCCCAAA	GGGGCAGTGT	TACCTTACTC	CTTCACTGCT	TCTTAGAAGG	TAGAATTAAG	2160
	TTTCTGGAAT	TGCACCTACA	TGTTTTCTTA	TTAACATTCA	GAATTTGGAA	TATTAATTTT	2220

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TCAGTGGAGT AGTTTTCGTA AATTGGTAAC TTGGAGAGTA AAATAACGTA TTTTGTCTTT 2280
CAATTGTTG TTGTTTACT TTTATGTAAA AATTGTATAT GTGAATTACA CAGTTCTAAT 2340
AAAACTCAT GCGTTTTCAT TACATCTAAT TTGAACCTCT AACTTCAGTG CCAGGAAGTGC 2400
TTTAAAGATG CTTTAAATGAA AAGTATTAG AAAATATATA GATTGTATG TCAGTTTATA 2460
CTTCAGAAAT CCATATATTT GTCATATTTA TTTTCTTAGA AACCTCCTAA TTGGATAACT 2520
AGATGGTATT TAAAAATGAAT GCCCAAAAAAT ATCTTGTAAC TTTGTCCAAA AGTTTATCTG 2580
TTGGAAGCCG CCAGCCATTTC ATGTAGAGAG TTTATAAGAA AATAATTTAA AATTGTATGC 2640
ATTTTATATT ACTATGGTAT CTGTGTACCA TATTCTTAAG TATTCAATTAT TAAATTGGTA 2700
CTTCTTAAAA CCATAACCTG GCTTGCTTT TAGTGTAAA CACAAATCC AACATTGTAT 2760
ATAGAGATTC TTCTTTTATG AAGAAGAGCT GACGTAATTT ATTACCACTG CATCTGCACA 2820
AAGACATTAA CATAAGTCTC TGAGCAGTGA TACATTTTCA AACATGAAGA GTGACACCA 2880
CCACATTAAA CAACCAACGGC AACACTCAGA CTTGGCACTT TCCTACGAAT CCATCCTATA 2940
TGTGCTGGT ATCGCTCTG GCATACTTA CACGAATCGT CCTCCTACT TGTCTACGCT 3000
CCTTCATCA GCACCTGCCA ACACATTCAC CTCTAACTTG TACAACCTTA CCAACTCACC 3060
ACAACATCTG CAACTCTACC CTATCAACTG CCAACCTAAA GACCCCAAC ACAACACAAC 3120
CCCCAACAC AAAACCACTA AATCATAACC ACCACACAG CCACACACCA CACACCCACC 3180
CACACAACCA ACACACACG ACCAAACACC CCACCAAAA CAAGCTAACA ACCCAAAACA 3240
GACAACACAT CACATACACT CACTACCCCC CCATACTCCC ACCCACCA

Seq ID NO: 234 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 27-281

1 11 21 31 41 51
AGCAGGAGGA GAGCTGGCGG GAAGACATGC ACCCCTTGAA GACCCAGAGA GAGGCGCTCT 60
GTCTACCGCG TAGCAGTTAC ATCAGACTGA GACACTTCCT GTTTACAGGA GACTATAAAA 120
TTCTGCTCCC GTGCTCATTT GGGGCTGACG CCATTTTAGG CCTCAGCCCA TCTGCACCCA 180
GGCGCTCACT GAAACAGTGT GTTGTCCAC ACCGCTTGT TTTGCTTGT GGCCTGCTCT 240
CAGGCTTCG ACCAATCCAA GAGCCTTGCA GAAAGCATTA ACGTGTCTTT CTCTTTGGCA 300
GAGTTTCTCT TTGCTCTGAT CTTGGAGACA TCCCTCTGCC TAGTGGAAAC ATAAGGAATA 360
CAGAAAGAA GCAGAGGAGAT AGACCAACGT GAGATTCTCC TTCTATGCACT CAAGAGAAAG 420
ATGTTGCAGG AAGAGCTAGT CTTTCAGGCT GGGCTGGTGA CCTGAGAAAG AATGTCCAGC 480
TTTTCTCTCT CACTTGGCAT ATCAAGAGCC AGGCGTGGAA GACTAAAACA GGAAATGTTT 540
ATAAAAACTG TTCAGCGGTT CGCCAACAAG AAGTGGTAAA GTAGCAAAAA TGGGGATGGA 600
GATGCCAGGA GGAAAGATGC CAGGGGTAAA GTGGGAAAT GGGAACTGGA AGCCAGGAGG 660
TCAAGCCAA GCACACAGTG TTCTGTTTT CATCACAGAA CTAATAAGTG GTGCTGAGGA 720
CTCAAAACCG GGAAGCCCA CTCTAGAAC CATGCTGCTC ATCCATATCC CCAAGGCCCT 780
GGTCAGAA CAAGCTAAGCA GATGGCTTGG GTCATCAGGA CGTCCATTAC ATCCAAAGGA 840
AGACAGCCTG TGACGTTTCA AAGCAAAAAG TCCCTACCA GCCAGTGAAG CTACCTGATT 900
TCTCAGTATC TTACGCCAG TGACACGATC TACCCTCAA ACTTAAAAAA AAAAGGGGAA 960
CATAAACACA TAACAGCAGC AGCAATAATT AAAGATGAGA TGAGAACAAT TAAGAAAAAA 1020
GGAAAGGCTC CTGTGACTG TTTTATTTTT AGGGAACAG AGAGGAAGAA GAATGATTTT 1080
TCTTTTGATG ACTCTATATC CAACTCTGAG GTTTGATTAA AGAATGACC TTGAACCACA 1140
GCAAGAAAA ATAAAGACA ATTTCCAGTA AGTATGCCAG TTCGAATTAA TGATTACTT 1200
TTTATTTTAA AACTGAATTC AGCAGAGATT TACATGCATT ACGATGATTA ACATCTGAAA 1260
TTTGACCTTG AATAATCTT TACATGTAA ATTCTTAATG ATCAAAACAA GGTCTCTAGT 1320
GATTAAAA CA TATTAGTAAT TAATTATTAA AGGAGAATAA TTGCAATAC AACATTCTTA 1380
AAATCTCAAG GCTTTTAAAG CATTTGTACA AATGACTGGA CATTTTAA ATTTGAAAAA 1440
AAAAAAGAG CCTCCATCTG ATTCTCATTT TCATTGTCTG TGCAACAACA AAAAAGGTAT 1500
GCATCTCTCT TCTCATTTTC CACTGTCTCG CAAGCTAGAA ATTCTACGA CTACCTTTGA 1560
TCCCATCAA GCCAAGAAA GAAAAGAAA TTGTTCTGTA CAGATATATG ACATTAATAA 1620
ATAATCCC

Seq ID NO: 235 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MHLPLKQREA VCLPRSSYIR LRHLFTGDIY KIPAPCSFGA DAILGLSPSA PRRLKQCV 60
PHRLVLLVGA LSGFRPIQEP CRKH

Seq ID NO: 236 DNA sequence
Nucleic Acid Accession #: NM_002075
Coding sequence: 406..1428

1 11 21 31 41 51
CCACAATAGG GGCAGACCTG TCCATCCTTC TCTGTGGGTC CCTGTACCT TTCTCCCCCA 60
ACAGGATCAG ACCCAGAGGC AGCTGGTTGG GGTGTGTCGA GAAGAAGGAT TATCCAGATC 120
AGTCTTTCT ATCTCAGCT CCTGCTGTA CCTCCCATTA CTCACCAAC CCTCTTCCCC 180
ACCACCTTGA GCTGAGGAGC ACAGTTTGAG GCCCCCCCA CCCCCCGCG GTGCGGGCCA 240
GGCCAGGCCA GGCAGCTCC TCTGGCAGCA GAGCCTGGGC AGGTGACGGG CGGCGCGGG 300
CGTCGCACT GAGGAGTAA GAGGCTCCC AGGAACCGGA GCTGGAACCC CGGCGAGGT 360
CCAGCCAGAG CCCAAGAGCC AGAGTGACCC CTCGACCTGT CAGCCATGGG GGAGATGGAG 420
CAACTGCTG AGGAAGCGGA GCAGCTCAAG AAGCAGATTG CAGATGCCAG GAAAGCCTGT 480
GCTGACGTTA CTCTGGCAGA GCTGGTGTCT GGCTAGAGG TGGTGGGACG AGTCCAGATG 540
CGGACGCGGC TGACGTTAAG GGGACACCTG GCCAAGATTT ACGCCATGCA CTGGGCCACT 600
GATTCTAAGC TGCTGTGAAG TGCCTGCAA GATGGGAAGC TGATCGTGTG GGACAGCTAC 660
ACCACCAACA AGGTGCACGC CATCCACTG CGCTCCTCT GGGTCATGAC CTGTGCTTAT 720
GCCCATCATG GGAATTTGT GGCATGTGG GGGCTGGACA ACATGTGTTT CATCTACAAC 780
CTCAAAATCCC CTGAGGGCAA TGTCAAGGTC AGCCGGGAGC TTTCTGCTCA CACAGTTAT 840
CTCTCTGCT CCGCTTCTT GGATGACAAC AATATTGTGA CAGCTCGGG GGACACCAAG 900
TGTGCTCTGT GGGACATTGA GACTGGCAG CAGAAGACTG TATTGTGGG ACACACGGGT 960
GACTGCATGA GCTGGCTGT GTCTCTGAC TTCAATCTCT TCATTTCGGG GGCCTGTGAT 1020
GCCAGTGCCA AGCTCTGGGA TGTGCGAGAG GGGACCTGCC GTCAGACTTT CACTGGCCAC 1080

5
10
15

```

GAGTGGGACA TCAACGCCAT CTGTTTCTTC CCCAATGGAG AGGCCATCTG CACGGGCTCG 1140
GATGAGCCTT CCGCGCGCTT GTTTGACCTG CGGGCAGACC AGGAGCTGAT CTGCTTCTCC 1200
CACGAGAGCA TCATCTGCGG CATCAGCTCC GTGGCCTTCT CCCTCAGTGG CCGCTACTA 1260
TTGCTGGCTT ACACGCACTT CAACTGCAAT GTCTGGGACT CCATGAAGTC TGAGCGTGTG 1320
GGCATCTCTT CTGGCCACGA TAACAGGGTG AGCTGCCTGG GAGTCACAGC TGACGGGATG 1380
GCTGTGGCCA CAGGTTCTCT GGACAGCTTC CTCAAAATCT GGAAGTGAAG AGGCTGGAGA 1440
AAGGGAAGTG GAAGGAGTG AACACACTCA GCAGCCCCCT GCCCGACCCC ATCTCATTCA 1500
GGTGTCTCTT TCTATATTCC GGGTGCCATT CCCACTAAGC TTTCTCTTTT GAGGGCAGTG 1560
GGGAGCATGG GACTGTGCCTT TTGGGAGGCA GCATCAGGGA CACAGGGGCA AAGAACTGCC 1620
CCATCTCTCT CCATGGCCTT CCCTCCCCAC AGTCTCTACA GCCTCTCCCT TAATGAGCAA 1680
GGACAACTGG CCCCTCCCCA GCCCTTTGCA GGCCAGCAG ACTTGAGTCT GAGGCCCCAG 1740
GCCCTAGGAT TCCTCCCCCA GAGCCACTAC CTTTGTCCAG GCCTGGGTGG TATAGGGCGT 1800
TTGGCCCTGT GACTATGGCT CTGGCACCAC TAGGGTCTCT GCCCTCTCTT TATTCTATGT 1860
TTCTCTCTTT TCTACCTTTT TTTCTCTCTT AAGACACCTG CAATAAAGTG TAGCACCCCTG 1920
GT

```

Seq ID NO: 237 Protein sequence:
Protein Accession #: NP_002066

20
25

```

1 11 21 31 41 51
MGEMEQLRQE AEQLKKQIAD ARKACADVTL AELVSGLEV GRVQMRTRT LRGLHLAKIYA 60
MHWATDSKLL VSAQSGGKLI VMSYTTNKV HAIPLRSSW MTCAYAPSGN FVACGGILDNM 120
CSINYLNKRE GNVKVSRELS AHTGYLSCCR FLDDNNIVTS SGDTTCALWD IETGQQTQVF 180
VGHYTDGMSL AVSPDFNLFI SGACDASAKL WDVREGTCRQ TFTGHESDIN AICPPFPNGEA 240
ICTGSDDASC RLFLRLADQE LICFSHESII CGITSVAFSL SGRLLFAGYD DFNCNVWDSM 300
KSERVIGILSG HDNRVSLCLGV TADGMAVATG SWDSFLKIWN

```

Seq ID NO: 238 DNA sequence
Nucleic Acid Accession #: CAT cluster

30
35
40
45
50

```

1 11 21 31 41 51
TCCCAATGTG TNGAACCTAC CATAAATTCT TTTCTTACNG GACAATCTTA TNCTAANCAA 60
TACCATTGCT TTTTAAGGCA GATAATCCTC CAAGTTTCTT AATGATATCT GAAACTATTA 120
ACTGATTCTG TGAATTATGA AATCTGAAAA GGAATTGGAA GTTGCTAAAA ATCTATCATT 180
TGCAATTGACC AGTGTGAAGC ACAGTGGAAAT GAGAATGGCT GCCCTGACAC CAAAGAAAAA 240
TAAGTGACTG GAAAGCTGAA GAATCACCGG CTTCAGTGAC ATGGAACCCA GTGATTGAT 300
TTTGTAGCAG TATCGGGTGA CTTTGAGGTG GTCAAGAAAC CACACTTTAA GAACAATGTC 360
CAAAAAGGGG AAAAAAAGA GCAACCAAAG AAAAAAATC CATAAAAATG CACAGAAGAA 420
AAGAAAGAAA AATAAAATAC ACAATATGGA CGATGGAGAA AAACAGTTAC ATTTCTTTAT 480
GGATCAAGAA GTTTGTGTAC ACATAATCTC ATTTTGAGAT ATATAACTAT TTTTGTCTTT 540
CAGAAAGTGA TCAAAATATT TCAAAATGCT GTCTTATGAA ACTACAATAT TCTCACAGAT 600
TAGAAAGTCT TTTCTGTAAA AGTCAGATAG TAAATATTTT AGGTTTGTCA GTGCTCTTTG 660
CAACTACTCA ACTTTCTCTC TGTAACACAA GAGTAGCTGT GGTACTGTGC AAATAAATTG 720
CTTGTTTCTC AATAAAGCTT CATTTACAAA AACATGCCAT GGGCCATATT TGGCCTGTAC 780
ACTGTTGTTT GCCAAGTCTT AATATAGTTG CTTAGCAAGT ATTGTGAGCT ATTTGAGGAA 840
GACATGAAGT TFCATTGGGT TGCTAAAAAG TATGTAGAAA TTCAAGGAAA AATTAAAAAT 900
TAGGCTAAGT TATAATACAC TGTTTTAACA ATTTGTAATA GTAAAGAGAA TTTACAAATA 960
AAAAATCCAA ATAAAA

```

Seq ID NO: 239 DNA sequence
Nucleic Acid Accession #: NM_001786.1
Coding sequence: 130-1023

55
60
65
70
75
80

```

1 11 21 31 41 51
GGGGGGGGGG GGCACCTTGGC TTCAAAGCTG GCTCTTGGAA ATTGAGCGGA GAGCGACGG 60
GTTGTTGTAG CTGCGCGTGC GGCGCGCGCG GAATAATAAG CCGGATCTA CCATACCCAT 120
TGACTAACTA TGAAGATTA TACCAAAATA GAGAAAATTG GAGAAGGTAC CTATGGAGTT 180
GTGTATAAGG GTAGACACAA AACTACAGGT CAAGTGGTAG CCATGAAAAA AATCAGACTA 240
GAAAGTGAAG AGGAAGGGGT TCCTAGTACT GCAATTCGGG AAATTTCTCT ATTAAGGAA 300
CTCGTCACTC CAAATATAGT CAGTCTTCAG GATGTGCTTA TGCAGGATTC CAGGTTATAT 360
CTCATCTTTG AGTTTCTTTC CATGGATCTG AAGAAATACT TGGATTCTAT CCCTCCTGGT 420
CAGTACATGG ATTCTTCACT TGTTAAGAGT TATTTATACC AAATCCTACA GGGGATTGTG 480
TTTTGTCACT CTAGAAGAGT TCTTCACAGA GACTTAAATC CTCAAAATCT CTGATTGAT 540
GACAAAGGAA CAATTAAACT GGCTGATTTT GGCTTGCCA GAGCTTTTGG AATACCTATC 600
AGAGTATATA CACATGAGGT AGTAACACTC TGGTACAGAT CTCAGAAGT ATTGCTGGGG 660
TCAGCTCGTT ACTCAACTCC AGTTGACATT TGGAGTATAG GCACCATATT TGCTGAACCTA 720
GCAACTAAGA AACCACTTTT CCATGGGGAT TCAGAAATTG ATCAACTCTT CAGGATTTTC 780
AGAGCTTTGG GCACTCCCAA TAATGAAGTG TGGCCAGAGG TGAATCTTTT ACAGGACTAT 840
AAGAATACAT TTCCCAAATG GAAACCAGGA AGCCTAGCAT CCCATGTCAA AAACCTGGAT 900
GAAAATGGCT TGGATTGCT CTOGAAAATG TTAATCTATG ATCCAGCCAA ACGAATTTCT 960
GGCAAAATGG CACTGAATCA TCCATATTTT AATGATTGG ACAATCAGAT TAAGAAGATG 1020
TAGCTTTCTG ACAAAGAGTT TCCATATGTT ATGTCAACAG ATAGTTGTGT TTTTATTGTT 1080
AACTCTGTGC TATTTTGTGC TTATATATAT TTCTTTGTTA TCAAACCTCA GCTGTACTTC 1140
GTCTTCTAAT TTCAAATAA TAACTTAAAA ATGTAAATAT TCTATATGAA TTTAAATATA 1200
ATTCTGTAAA TGTGAAAAA AAAAAAATA AAAAA

```

Seq ID NO: 240 Protein sequence:
Protein Accession #: NP_001777.1

85

```

1 11 21 31 41 51
MEDYTKIEKI GEGTYGVYK GRHKTGQVY AMKKIRLESE EBGVPSTAIR EISLLKELRH 60
PNIVSLQDVL MQDSRLYLIF EFLSMDLKKY LDSIPPQYQ DSSLVKSILY QILQGVTFCH 120

```


SRRLHRLDK PQLNLIDDKG TIKLADFLGA RFGIPIRVY THEVVTLWYR SPEVLLGSAR 180
 YSTPVDWNSI GTIFAEALATK KPLFHGDSEI DQLFRIFRAL GTFNNEVWPE VESLDQYKNT 240
 FPKMKPGSLA SHVKNLDENG LDLLSKMLIY DPAKRISGRM ALNHPYPNDL DNQIKRM

5

Seq ID NO: 241 DNA sequence
 Nucleic Acid Accession #: NM_033379.1
 Coding sequence: 132-854

10

1 11 21 31 41 51
 CGCCCGCGCG CGGGCTCAAC TTGTAGAGC GAGGGGCCAA CTGGCAGAG CGCGCGGCCA 60
 GCTTTGCAGA GAGCGCCCTC CAGGGACTAT GCGTGGCGGG ACACGGGATC TACCCATACC 120
 ATTGACTAAC TATGGAAGAT TATACCAAAA TAGAGAAAAT TGGAGAAGST ACCTATGGAG 180
 TTGTGTATAA GGGTAGACAC AAAACTACAG GTCAAGTGGT AGCCATGAAA AAAATCAGAG 240
 TAGAAAGTGA AGAGGAAGGG GTTCTTAGTA CTGCAATTCC GGAATTTCTT CTATTAAAGG 300
 AACTTCGTGA TCCAAATATA GTCACTCTTC AGGATGTGCT TATGCAGGAT TCCAGGTTAT 360
 ATCTCATCTT TGAGTTCTCT TCCATGGATC TGAAGAAATA CTGGATTCTT ATCCCTCCTG 420
 GTCAGTACAT GGATTCCTCA CTGTTAAGG TAGTAACACT CTGGTACAGA TCTCCAGAAG 480
 TATTGCTGGG GTCAGCTCGT TACTCACTC CAGTTGACAT TTGGAGTATA GGCACCATAT 540
 TTGCTGAATC AGCAACTAAG AAACCACTTT TCCATGGGGA TTCAGAAATT GATCAACTCT 600
 TCAGGATTTT CAGAGCTTTG GGCACCTCCA ATAATGAAGT GTGGCCAGAA GTGGAATCTT 660
 TACAGGACTA TAAGAATACA TTTCCCAAT GGAACCCAGG AAGCCTAGCA TCCCATGTCA 720
 AAAACTTGGG TGAATATGGC TTGGATTTCG TCTCGAAAAT GTTAATCTAT GATCCAGCCA 780
 AACGAATTTC TGGCAAAATG GCACTGAATC ATCCATATTT TAATGATTGG GACAATCAGA 840
 TTAAGAAGAT GTAGCTTTCT GACAAAAAGT TTCCATATGT TATGTCAACA GATAGTTGTG 900
 TTTTATATGT TAACCTCTGT CTATTTTGT CTTATATATA TTTCTTTGT ATCAAACTTC 960
 AGCTGTACTT CGTCTTCTAA TTTCAAAAAT ATAACCTAAA AATGTAATA TTCTATATGA 1020
 ATTTAAATAT AATTCGTGTA ATGTGAAAAA AAAAAAAA AAAAAA

30

Seq ID NO: 242 Protein sequence:
 Protein Accession #: NP_203698.1

35

1 11 21 31 41 51
 MEDYTKIEKI GEGTYGVVYK GRHKTGTQVQV AMKKIRLESE EEGVPSTAIR EISLLKELRH 60
 PNIVSLQDVL MQDSRLYLIF EFLSMDLKKY LDSIPFGQYM DSSLVKVVTL WYRSPEVLLG 120
 SARYSTPVDI WSIGTIFAEL ATKPLPFHGD SEIDQLFRIP RALGTFNNEV WPEVESLDQY 180
 KMTFFPKMKPG SLASHVKNLD ENGLDLSKML LIYDPAKRIS GRMALNHPYP NDLDNQIKRM

40

Seq ID NO: 243 DNA sequence
 Nucleic Acid Accession #: AF101051.1
 Coding sequence: 221-856

45

1 11 21 31 41 51
 GAGCAACCTC AGCTTCTAGT ATCCAGACTC CAGCGCCGCC CGGGCGCGCG ACCCCAACCC 60
 CGAGCCAGAG CTTCTCCAGC GCGCGGCGCAG CGAGCAGGGC TCCCGCGCTT AACTTCTCTC 120
 CGGGGGGCCA GGCACCTTCG GGAGTCCCGG TTGCCCACTT GCAAACTCTC CGCTTCTGCG 180
 AACTGCCACC CCTGAGCCAG CGCGGGCGGCC CGAGCGAGTC ATGGCCCAACG CGGGGCTGCA 240
 GCTGTGTGGC TTCATTCTCG CCTTCTCTGG ATGGATCGGC GCCATGCTCA GCACTGCCCT 300
 GCGCCAGTGG AGGATTACTT CCTATGCCGG CGACAACATC GTGACCGCCC AGGCCATGTA 360
 CGAGGGGCTC TGGATGTCTT GCGTGTCCGA GAGCACCGGG CAGATCCAGT CCAAGTCTTT 420
 TGACTCTCTG CTGAATCTGA CGACACATT GCAAGCAACC CGTGCCTTGA TGGTGGTTGG 480
 CATCTCTCTG GGAGTGATAG CAATCTTTGT GGCCACCGTT GGCATGAAGT GTATGAAGTG 540
 CTTGGAAGAC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTTCT 600
 TCTTGCAAGT CTGGCTATTT TAGTTGCCAC AGCATGGTAT GGCATAGAA TCGTTCAAGA 660
 ATTCTATGAC CTTATGACCC CAGTCAATGC CAGGTACGAA TTTGGTCAGG CTCTCTTCC 720
 TGGCTGGGCT GCTGCTCTC TCTGCTTCTT GGGAGGTGCC CTACTTTGCT GTTCTCTGTC 780
 CGGAAAAACA ACCTCTTACC CAACACCAAG GCCCTATCCA AAACCTGCAC CTTCCAGCGG 840
 GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAAAATCA TGTGAAACA AACCGAAAAT 900
 GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GTAATCTGAA 960
 GTATGGTATT ACAAAACAAA CAAAACCAAT AAAAAACCAT GTGTAAATAT ACTCAGTGCT 1020
 AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080
 TTGTATTACT GCTTCCCATT GAGTAATCAT ACTCAAATGG GGGAGGGGGT GCTCCTTAAA 1140
 TATATATAGA TATGTATATA TACATGTTT TCTATTAAAA ATAGACAGTA AAATACTATT 1200
 CTCATTATGT TGATACTAGC ATACTTAAAA TATCTCTAAA ATAGGTAAAT GTATTTAATT 1260
 CCATATTGAT GAAGATGTTT ATTGGTATAT TTTCTTTTTC GTCTTATAT ACATATGTAA 1320
 CAGTCAATA TCATTACTC TTCTTCTTGA GCTTTGGGTG CCTTTGCCAC AAGACCTAGC 1380
 CTAATTTACC AAGGATGAAT TCTTTCAATT CTTATGCGT GCCCTTTTCA TATACTTATT 1440
 TTATTTTTTA CCATAATCTT ATAGCACTTG CATCGTTATT AAGCCCTTAT TTGTTTGTG 1500
 TTTCAATGGT CTCTATCTCC TGAATCTAAC ACATTTCATA GCCTACATT TAGTTTCTAA 1560
 AGCCAAGAAG AATTTATTAC AAATCAGAAC TTTGGAGGCA AATCTTTCTG CATGACCAA 1620
 GTGATAAATT CCGTTGACC TTCCACACA ATCCCTGTAC TCTGACCCAT AGCACTCTTG 1680
 TTTGCTTTGA AAATATTGTT CCAATTGAGT AGCTGCATGC TGTCCCCCA GGTGTTGTAA 1740
 CACAACPTTA TTGATTGAAT TTTTAAGCTA CTTATTCTAT GTTTTATATC CCCCTAAACT 1800
 ACCTTTTGTG TCCCATTTCC TTAATTGTAT TGTTTCCCA AGTGAATTA TCATGGTTT 1860
 TATATCTTCC TAATAGGTG TGGTCTGTTT GTCTGAACAA AGTGCTAGAC TTTCTGGAGT 1920
 GATAATCTGG TGACAAATAT TCTCTCTGTA GCTGTAAGCA AGTCACTTAA TCTTCTTACC 1980
 TCTTTTCTCT ATCTGCCAAA TTGAGATAAT GATACTTAAC CAGTTAGAAG AGGTAGTGTG 2040
 AATATTAAAT AGTTTATATT ACTCTCATTC TTTGAACATG AACTATGCCT ATGTAGTGTG 2100
 TTTATTGTCT CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACCT ACACACGTAC 2160
 CTTATGTGTA TTCACCTGCT TCCTCTCTCT ACCAGTCTAT TTCCACTGAA CAAAACCTAC 2220
 ACACATACCT TCATGTGGTT CAGTGCCTTC CTCTCTCTAC CAGTCTATTT CCACTGAACA 2280
 AAACCTACGC ACATACCTTC ATGTGGCTCA GTGCCCTTCT CTCTCTACCA GTCTATTTC 2340
 ATTCTTTCAG CTGTGCTGTA CATGTTTGTG CTCTGTTCCT TTTTAAACAC TGCTCTTACT 2400
 TTTCCAGTCT GTACAGAATG CTATTTCTACT TGAGCAAGAT GATGTATGGA AAGGGTGTG 2460

85

5 GCACCTGGTGT CTGGAGACCT GGATTGAGT CTGGTGCTA TCAATCACCG TCTGTGTTTG 2520
AGCAAGGCAT TTGGCTGCTG TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAATTC 2580
CTGATCTTCC CACCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640
GTGGTTTGTG AATTGAAAA GTGCTATACT AAGGAAAAA ATTGAGGAAT TAACTGCATA 2700
CGTTTGGTGT TTGCTTTTCA AATGTTTGAA AATAAAAAA TGTAAAGAAA TGGGTTTCTT 2760
GCCTTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAACCGAAT 2820
AAGATTCTGA GGAAGTCTTA TCCTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880
ACAGATGTAA TGGGAAGAAA TAAAGCCTA CGTGTGTTG AATCCAACAG CAAGGGAGAT 2940
TTTGTAAATCA TAATAACTCA TAAGGTGCTA TCTGTTCACT GATGCCCTCA GAGCTCTTGC 3000
10 TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGGTACT TCATAATAAA 3060
CTACACAAGG AAGTCAAGC ACCGTGCTTT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120
TGCTTCCAA ACCTGAGAAT ATATGCTTTT GGAAGTTAAA ATTTAAATGG CTTTGGCCAC 3180
ATACATAGAT CTTCAATGAT TGTGAGTGTA ATTCATATGT GATATCAGTT ACCAAACATT 3240
15 ACAAAAAAT TTTATGGCCC AAAATGACCA ACGAAATTGT TACAATAGAA TTTATCCAAT 3300
TTTGATCTTT TTATATTCTT CTACCACACC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360
TTATAATGGG AATTGTGATA AAGCATTACT CTTTTCAT AAATGTTTT TTAATTTAAA 3420
AAAAGGAAAA AAAAAAAA AAA

20 Seq ID NO: 244 Protein sequence:
Protein Accession #: AAD16433.1

25 1 11 21 31 41 51
MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMYEGE WMSCVSQSTG 60
QIQCKVFDL LNLSSLTQAT RALMVVGILL GVIAIPVATV GMKCMKCLEDEVQKMRMAV 120
IGGAIFLLAG LAILVATAWY GNRIVQEFYD PMTPVNARYE FGQALFTGWA AASLCLLGGA 180
LLCCSCPRKT TSYPTRPRYP KPAPSSGKDY V

30 Seq ID NO: 245 DNA sequence
Nucleic Acid Accession #: CAT cluster

35 1 11 21 31 41 51
TTTTTTTTT TTTTTTTTT TTTTCAAGG AGAGCACAAG GAACTTTATT AATGACTTTC 60
TTAATGGTTA AATGCTGTTT ACCAAGTGAC CCAGAGGCAG CGTGGTTTAG TGGTTTCAAC 120
AGCATGGTCC CGAGAGTCTG ACCAACCTCA GTTCAATCC TTCCTTTGTC TTCACTTAGT 180
TTTTCTTCCT GAGATTAGT TTCTTCATCG TTAACAATGA GGATATTAAAT ATGTTTCACA 240
40 CAGTTGTTAT GAAGAATGCA TATATTAGAA TGCTGTAGT CTCAGCTACT CAGGAGGCTA 300
AGGTGGGAG GTGCTCAAG CCCAGGAATT CAAAGCTGCA ATGCATTATG ATTACAGCTG 360
TTAATAGCCA CTGCACCTCA GCCTGGGCAA TGTAGTAAGA TCCCATCTCT GGCTCGGAGG 420
GTCTACGCC CACGAGTCT CGCTGATTGC TAGCACAGCA GTCTGAGATC AAATGCA

45 Seq ID NO: 246 DNA sequence
Nucleic Acid Accession #: XM_058553.2
Coding sequence: 897-1400

50 1 11 21 31 41 51
AATTTTCAGA AGTTTCGTAT GGGGATGGTT TTATATAAAT TCAGGTTTTT CCCACAATAA 60
TAAATGTATT TAGTCTCAGT GCTCAATAGA AGAGATTCTT AATAGAAAAG GATTCAAAC 120
GTGAACCAT TTCTCTTTA ATGTTTCACA TTCCGTGTAC AGATTGTGTC TCTGTGACT 180
CTGTTATCCA TAATATGGAC AGTTCTTGAG TCCTAACATT GAGAGGTTTT CCCTTAGTGC 240
ATAGAGGGAA TGAGATTAA TTGAGAGAAGC TTAAGTATT GCCACTTTAG CACTGAAGAT 300
55 TGGGATGAGA GGAGGTGAAA CCTCACTAGA AAAAGGGACA ATGTTAGTGT GGCCCTTCCT 360
GATCATGTTT AAGAAAGTCT ATGAAAATGG TGAAGTAGTG TTTCCAAGCA TATTGGAAGG 420
GTTGAGTGA TACTGTCTGT CAAAGACTTC CAGCATTTCC AGGTCTAGA GAGGAACAAG 480
ACTGGTAACC TGCTATCTG TATTTTAAAG AACCAGGAG GAAAGCTTTA TAATAGAACA 540
60 TTAATTTCTG GTTTATGTAT AAGGGGTTTT TTGTTTTT AAAGACAGGA TCTCACTCCA 600
TTGTCCAGCG CAGTGCAAT GGCACGAACC TCATAGCTCC TGGACTTAAG TGATCTGCCT 660
GCCTTTGCCT CCGAGTAGC TGGGACTACA GGCATGAGCC CCCATGCCTG GCTAAGTTTG 720
TTTTTTTGTG TGTGTTTGTG TTTGTTTTG GGGGGGGTTG TTTGTTTTT TGTAGAGACG 780
TAGTCTTGCT TTGTTGCCAG GCTAGTCTCA AACTCCTGGC TTCAGTGAT CCTCCTGCCT 840
CAGCCTCCCA GAGTGCTAGG ATTACAGCAC TTGGATTTCAG CTCTCTCAT TCCAACATGG 900
65 AAGAACTTA CACCGACTCC CTGGACCCCTG AGAAGCTATT GCAATGCCCC TATGACAAAA 960
ACCATCAAT CAGGCTTGC AGGTTTCTT ATCATCTTAT CAAGTGACAG AAGATCATC 1020
CTGATGTTGC AAGCAATG GCTACTGTG CTTCAATGC TCSCCACCAG GTTCCTCGAG 1080
CTGAATTTAG TCATCATATC TCAAGCTGTG ATGACAGAAG TTGTATTGAG CAAGATGTTG 1140
70 TCAACCAAC CAGGAGCCTT AGACAAGAGA CTCTGGCTGA GAGCACTGG CAGTGCCTC 1200
CTTGCGATGA AGACTGGGAT AAAGATTGT GGGAGCAGAC CAGCACCCCA TTTGTCTGGG 1260
GCACACTCA CTACTCTGAC AACACAGCC CTGCGAGCAA CATAGTTACA GAACATAAGA 1320
ATAACCTGGC TTCAGGCATG CGAGTTCCCA AATCTCTGCC GTATGTTCTG CCATGGAAAA 1380
ACAAATGAAA TGCACAGTAA CTGAATACCT ATCTCATCAA ATGCCAGACC CTAGAAGACT 1440
75 GTTGCTTCTT CTTCTACCA TGGGTTCTCA TTTTCTCTCT AATCTAATTA TAGAATGGTA 1500
AACTCCCTGT GACTTTCCTA ACTGACAAGC ACACTTTTT CCTCCCCCT TGAATCCTCA 1560
TTTAATGCAA GAACCTCAT ACTCAGAAGC TTCCAAATAA ACCTTGATA CAGATTG

80 Seq ID NO: 247 Protein sequence:
Protein Accession #: XP_058553.1

85 1 11 21 31 41 51
MEETYSLSL PEKLLQCFYD KNHQRACRF PYHLIKCRKN HPDVASKLAT CFPNARHQVP 60
RAEISRHIS CDRSCIEQD VVNQTRSLRQ ETLAEETWQC PPCDEDDWDK LNEQTSTFPV 120
WGTHYSIDNN SPASNIVTEH KNNLASGMRV PKSLPYVLPW KNNGNAQ

Seq ID NO: 248 DNA sequence
Nucleic Acid Accession #: NM_003392
Coding sequence: 758..1855

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
TTAAGGAAAT	CGGGCTGCT	CTTCCCCATC	TGGAAGTGGC	TTTCCCCACA	TGCGCTCGTA	60
AACGTATTAT	GAACATACG	ATGTTAATTC	GGAGCTGCAT	TTCCAGCTG	GGCACTCTCG	120
CGCGCTGGTC	CCCCGGGCTT	CGCCCCCCAC	CCCTGCGCCT	TCCTCCCGC	GTCTGCCCC	180
CATCTCTCAC	CCCCGGGCTT	GGCCACCCCG	CCTCTTGGC	AGCCTCTGCG	GGCAGCGCGC	240
TCCACTCGCT	TCCCGTGTCT	CTCTCGGCCA	TGGAATTAAT	TCTGGCTCCA	CTTGTGCTC	300
GGCCAGGTT	GGGAGAGGA	CGGAGGGTGG	CCGACGCGGG	TTCTGTAGTG	AATTACCCAG	360
GAGGCACTGA	GCACAGCAC	AAC TAGAGAG	GGT CAGGGG	GTGCGGACT	CGAGCGAGCA	420
GGAGGAGGCG	AGCGCTGGC	ACCAGGGCTT	TGACTCAACA	GAATTGAGAC	ACGTTTGTAA	480
TGCGTGGCGT	GGCCCGCGCA	CAGGATCCCA	GGGAAATCA	GATTTCCTGG	TGAGGTTGCG	540
TGGGTGGATT	AATTTGGAAA	AAGAAACTGC	CTATATCTTG	CCATCAAAAA	ACTCAGCGAG	600
GAGAAGCGCA	GTCAATCAAC	AGTAAACTTA	AGAGACCCCC	GATGCTCCCC	TGGTTAACT	660
TGTATGCTTG	AAAATTATCT	GAGAGGGAAT	AAACATCTTT	TCCTTCTTCC	CTCTCCAGAA	720
GTCCATTGGA	ATATTAGGCC	CAGGAGTTGC	TTTGGGGATG	GCTGGAAGTG	CAATGTCTTC	780
CAAGTCTTCT	CTAGTGGCTT	TGCGCATATT	TTTCTCTTTC	GGCCAGGTTG	TAATTGAAGC	840
CAATTCTTGG	TGGTGGCTAG	GTATGAATAA	CCCTGTTTCT	ATGTGAGAG	TATATATTAT	900
AGGAGCAGAG	CTCTCTGCA	GCCAACTGGC	AGGACTTCTT	CAAGGACAGA	AGAAACTGTG	960
CCACTTGTAT	CAGGACCACA	TGCAGTACAT	CGGAGAAGGC	GCGAAGACAG	GCATCAAGA	1020
ATGCCAGTAT	CAATTCCGAG	ATCGACGGTG	GAAGTGCAGC	ACTGTGGATA	ACACCTCTGT	1080
TTTTGGCAGG	GTGATGCAGA	TAGGCAAGCG	CGAGACGGCC	TTCACTACAG	CCGTGAGCGC	1140
AGCAGGGGTG	GTGAAGCCCA	TGAGCGGGCC	GTGCGCGGAG	GGCGAGCTGT	CCACCTGCGG	1200
CTGACGCGCG	GGCGCGCGCC	CCAAGGAOCT	GGCGGGGAC	TGGCTCTGGG	GGCGCTGCGG	1260
CGACAACATC	GACTATGGCT	ACCGCTTTGC	CAAGGAGTTC	GTGGACGCCC	GGAGCGGGA	1320
GGCATCTCAC	GCCAAAGGCT	CCTACGAGAG	TGCTCGCATC	CTCATGAACC	TGCACAACAA	1380
CGAGGCGCGC	CGCAGGACGG	TGTACAACCT	GGCTGATGTG	GCCTGCAAGT	GGCATGGGGT	1440
GTCCGGCTCA	TGTAGCCTGA	AGACATGCTG	GCTGCAGCTG	GCAGACTTCC	GCAAGGTGGG	1500
TGATGCCCTG	AAGGAGAAGT	ACGACAGCGC	GGCGGCCATG	CGGCTCAACA	GGCGGGGCAA	1560
GTGTGTACAG	GTCAACAGCC	GCTTCAACTC	GGCCACCACA	CAAGACCTGG	TCTACATCGA	1620
CCCCAGCCCT	GACTACTGCG	TGCGCAATGA	GAGCACCAGC	TCGCTGGGCA	CGCAGGGCGG	1680
CCTGTGCAAC	AAGAGCTCGG	AGGSCATGGA	TGGCTGCGAG	CTCATGTGCT	CGCGCGGTGG	1740
GTACGACCAAG	TTCAGACCGC	TGCAGACGGA	GCCTGCCAC	TGCAAGTTCC	ACTGGTGGCT	1800
CTACGTCAAG	TGCAAGAAGT	GCACGAGAT	CGTGGACCAG	TTGTGTGCA	AGTAGTGGGT	1860
GGCACCAGAG	ACTCAGCCCC	GCTCCAGGTA	CCGCTTATT	TATAGAAAGT	ACAGTGATT	1920
TGGTTTTTGG	TTTTTAGAAA	TATTTTTTAT	TTTTCCCCAA	GAATTGCAAC	CGGAACCAAT	1980
TTTTTCTCTG	TTACCATCTA	AGAACTCTGT	GGTTTATTAT	TAATATTATA	ATTATTATTT	2040
GGCAATAATG	GGGGTGGGAA	CCACGAAAAA	TATTTATTTT	GTGGATCTTT	GAAGAAGTAA	2100
TACAAGACTT	CTTTTGGATA	GTATAGAATG	AAGGGGGA	TAACACATAC	CCTAACTTAG	2160
CTGTGTGGGA	CATGGTACAC	ATCCAGAAGG	TAAAGAAATA	CATTTCTTTT	TTCTCAATA	2220
TGCCATCTATA	TGGGATGGGT	AGGTTCCAGT	TGAAAGAGGG	TGGTAGAAAT	CTATTACAA	2280
TTCACTCTCT	ATGACCAAAA	TGAGTTGTAA	ATTTCTCTGT	GCAAGATAAA	AGGCTCTGGG	2340
AAAACAAAAC	AAAACAAAAC	AAACCTCCCT	TCCCCAGCAG	GGCTGCTAGC	TTGCTTTCTG	2400
CATTTTCAAA	ATGATAATTT	ACAATGGAAG	GACAAGAATG	TCATATTCTC	AAGGAAAAAA	2460
GGTATATCAC	ATGCTCTCAT	CTCCTCAAT	ATTCCATTG	CAGACAGACC	GTCAATTCT	2520
AATAGCTCAT	GAAATTTGGG	CAGCAGGGAG	GAAAGTCCCC	AGAAATTA	AAATTTAAAA	2580
CTCTTATGTC	AAGATGTGGA	TTTGAAGCTG	TTATAAGAAT	TGGGATTCCA	GATTGTGAAA	2640
AAGACCCCCA	ATGATTCTGG	ACACTAGATT	TTTGTGTTGG	GGAGGTTGGC	TTGAACATA	2700
ATGAATATTC	CTGTATTTTC	TTAGGGATAC	TTGGTTAGTA	AAATATAATA	GTAGAAATA	2760
TACATGAATC	CCATTACAGC	GTTTCTCAGC	CCAGCAACA	AGGTAAATG	GTGCCATTCA	2820
GCACTGCACC	AGAGCAGACA	AOCTATTGTA	GGAAAAACAG	TGAAATCCAC	CTTCTCTTTC	2880
ACACTGAGCC	CTCTCTGATT	CCTCCGTGTT	GTGATGTGAT	GCTGGCCACG	TTTCCAAAGC	2940
GCAGCTCCAC	TGGGTCCCC	TTGGTTGTAG	GACAGGAAT	GAAACATTAG	GAGCTCTGCT	3000
TGGAAACAG	TTCACTACTT	AGGGATTTTT	GTTCCTTAAA	ACTTTTATTT	TGAGGAGCAG	3060
TAGTTTTCTA	TGTTTTAATG	ACAGAACTTG	GCTAATGGAA	TTACAGAGG	TGTTGCAGCG	3120
TATCACTGTT	ATGATCCTGT	GTTTAGATTA	TCCACTCATG	CTTCTCTAT	TGTAAGTCA	3180
GTGTACTCTA	AACTGTCTCC	CAGTGTACTT	GAACAGTTGC	ATTATATAAG	GGGGAATGT	3240
GGTTTAAATG	TGGCTGATAT	CTCAAGTCT	TTGTACATA	ACATATATAT	ATATATACAT	3300
ATATATAAAT	ATAAATAA	ATATATCTCA	TTGCAGCCAG	TGATTAGAT	TTACAGCTTA	3360
CTCTGGGGTT	ATCTCTCTGT	CTAGAGCAT	GTGTCTCTTC	ACTGCACTCC	AGTTGGGATT	3420
ATTCCAAAG	TTTTTTGAGT	CTTGAGCTTG	GGCTGTGGCC	CGCTGTGAT	CATACCTGTA	3480
GCACGACGAA	GCAACCTCGT	TTCTGAGGAA	GAAGCTTGAG	TTCTGACTCA	CTGAAATGCG	3540
TGTTGGGGTT	AAGATATCTT	TTTTTCTTTT	CTGCCTCACC	CCTTTGTCTC	CAACCTCCAT	3600
TTCTGTCTAC	TTTGTGGAGA	GGGCATTACT	TGTTCTGTTAT	AGACATGGAC	GTTAAGAGAT	3660
ATTCAAACT	CAGAAGCATC	AGCAATGTTT	CTCTTTTCTT	AGTTCAATTCT	GCAGAAATGGA	3720
AACCCATGCC	TATTAGAAAT	GACAGTACTT	ATTAATTGAG	TCCCTAAGGA	ATATTACAGC	3780
CACTACATAG	ATAGCTTTTT	TTTTTTTTTT	TTTTTTTTAA	TAAGGACACC	TCTTTCCAAA	3840
CAGGCCATCA	AATATGTTCT	TATCTCAGAC	TTACGTTGTT	TTAAAAGTTT	GGAAAGATAC	3900
ACATCTTTTC	ATACCCCCC	TTAGGAGGTT	GGGCTTTTCT	ATCACCTCAG	CCAAGTGTGG	3960
CTCTTAATTT	ATGATCAAT	GATATCCACA	TCAGCCCACT	GTGGCTCTTT	AATTATTGTC	4020
ATAATGATAT	TCACATCCCC	TCAGTTGCAG	TGAATTGTGA	GCAAAAGATC	TTGAAAGCAA	4080
AAAGCACTAA	TGAGTTTAAA	ATGTCACCTT	TTTGGTTTTT	ATTATACAAA	AACCATGAAG	4140
TACTTTTTTT	ATTGTCTAAA	TCAGATTGTT	CCTTTTTAGT	GACTCATGTT	TATGAAGAGA	4200
GTGTAGTTTA	ACAATCCTAG	CTTTTAAAAG	AAACTATTTA	ATGTAAAATA	TCTTACATGT	4260
CATTCTAGATA	TATGTATAT	CTTCTAGCCT	TTATCTGTGA	CTTTTAAATG	ACATATTCTT	4320
GTCTTGGCTG	ATTGTATAT	TTCACTGGTT	TAAAAACAA	ACATCGAAAG	GCTTATTCCA	4380
AATGGAAGAT	AGAATATAAA	ATAAACGTT	ACTTGTAAAA	AAAAAAA		

Seq ID NO: 249 Protein sequence:
Protein Accession #: NP_003383

85
1 11 21 31 41 51

MAGSAMSSKF FLVALAIFFS FAQVVEANS WWSLGMNFV QMSEVYIIGA QPLCSQLAGL 60
 SQGQKLCHEL YQDHMQYIGE GAKTGIKECQ YQFRHRRWNC STVDNTSVFG RVMQIGSRET 120
 APTYAVSAAG VVNAMSRACR EGELSTCGCS RAARPKDLPR DMLNGGCGDN IDYGYRFAKE 180
 FVDARERERI HAKGSYESAR ILMNLENNEA GRRTVYNLAD VACKCHGVSG SCSLKTCWLQ 240
 LADPRKVGDA LKEKYDSAAA MRLNSRGLV QVNSRFNSPT TQDLVYIDPS PDYCVRNEST 300
 GSLGTQGRLC NKTSBGMDCG ELMCCGRGYD QFKTVQTERC HCKFHWCCYV KCKKCTEIVD 360
 QFVCK

Seq ID NO: 250 DNA sequence
 Nucleic Acid Accession #: NM_014058
 Coding sequence: 56..1324

1 11 21 31 41 51
 TGACTTGGAT GTAGACCTCG ACCTTCACAG GACTCTTCAT TGCTGGTTGG CAATGATGTA 60
 TCGGCCAGAT GTGGTGAGGG CTAGGAAAAG AGTTTGTGG GAACCCCTGG TTATCGGCCT 120
 CGTCATCTTC ATATCCCTGA TTGTCTGGC AGTGTGCATT GGACTCACTG TTCATTATGT 180
 GAGATATAAT CAAAAGAAGA CCTACAATTA CTATAGCACA TTGTCTATTA CAACTGACAA 240
 ACTATATGCT GAGTTTGGCA GAGAGGCTTC TAACAATTTT ACAGAAATGA GCCAGAGACT 300
 TGAATCAATG GTGAAAAATG CATTTTTATA ATCTCCATTA AGGGAAGAAT TTGTCAAGTC 360
 TCAGGTTATC AAGTTCAGTC AACAGAAAGA TGGAGTGTGG GCTCATATGC TGTGTGTTTG 420
 TAGATTTTAC TCTACTGAGG ATCTCGAAAC TGTAGATAAA ATTGTTCAAC TTGTTTACA 480
 TGAAGAAGCTG CAAGATGCTG TAGGACCCCTC TAAAGTAGAT CCTCACTCAG TTAATAATTA 540
 AAAAATCAAC AAGACAGAAA CAGACAGCTA TCTAAACCAT TGCTGCGGAA CACGAAGAAG 600
 TAAACTCTTA GGTTCAGAGT TCAGGATCGT TGGTGGGACA GAAGTAGAAG AGGGTGAATG 660
 GGCCTGGGAG CGTAGCCTGC AGTGGGATGG GAGTCATCGC TGTGGAGCAA CCTTAATTA 720
 TGCCACATGG CTGTGTGAGT CTGCTCACTG TTTTACAACA TATAAGAACC CTGCCAGATG 780
 GACTGCTTCC TTGTGAGTAA CAATAAAACC TTGAAAATG AAACGGGGTC TCCGAGAAAT 840
 AATTGTCCAT GAAAAATACA AACACCCATC ACATGACTAT GATATTCTTC TTGCAGAGCT 900
 TTCTAGCCCT GTTCCCTACA CAAATGCAGT ACATAGAGTT TGCTCCCTG ATGCATCCTA 960
 TGAGTTTCAA CCAGGTGATG TGATGTTTGT GACAGGATTG GGAGCACTGA AAAATGATGG 1020
 TTACAGTCAA AATCATCTTC GACAAGCACA GGTGACTCTC ATAGACGCTA CAACCTGCAA 1080
 TGAACCTCAA GCTTACAATG ACGCCATAAC TCCTAGAATG TTATGTGCTG GCTCCTTAGA 1140
 AGGAAAAACA GATGCATGCC AGGGTGACTC TGGAGGACCA CTGGTTAGTT CAGATGCTAG 1200
 AGATATCTGG TACCTTGCTG GAATAGTGAG CTGGGAGATG GAATGTGCGA AACCCACAA 1260
 GCCTGGTGT TATACAGAGT TTACGGCCTT CGGGGACTGG ATTACTTCAA AAACCTGGTAT 1320
 CTAGAGAGA AAGGCTCAT GGAACAGATA ACATTTTITT TTGTTTITTG GGTGTGGAGG 1380
 CCATTTTATG AGATACAGAA TTGGAGAAGA CTGCAAAAC AGCTAGATTG GACTGATCTC 1440
 AATAAATCTG TTGCTTGATG CAAAAAATA A

Seq ID NO: 251 Protein sequence:
 Protein Accession #: NP_054777

1 11 21 31 41 51
 MYRPDVVRAR KRVCWEPWVI GLVIFISLIV LAVCIGLTVH YVRYNQKRTY NYYSTLSFTT 60
 DKLYAEFGRE ASNPTFMSQ RLESMVKNF YKSPLREEPV KSVIKFSSQ KGVLAEMLL 120
 ICRPHSTEDP ETVDKIVQLV LHEKLQDAVG PPKVDPHSVK IKKINKTETD SYLNHCOCGR 180
 RSKTLGQSLR IVGGTEVEEG EWPWQASLQW DGSHRGATL INATWLVSAA HCPFTYKNPA 240
 RWTASPGVTI KPSKMKRLR RIIVHEKYKH PSHDYDISLA ELSSFPVPTN AVERVCLPDA 300
 SYEFQPGDVM FVTGFGALIN DGYSQNLHQ AQTLLIDATT CNEPQAYNDA ITPRMLCAGS 360
 LEGKTDACQG DSGGPLVSSD ARDIWYLAGI VSWGDECAKP NKPGVYTRVT ALRDWITSKT 420
 GI

Seq ID NO: 252 DNA sequence
 Nucleic Acid Accession #: NM_003504.2
 Coding sequence: 71-1771

1 11 21 31 41 51
 GGCACGAGGC CTGCTGCCGC CGGGCTCTTG GTACCTCAGC GCGAGCGCCA GCGCTCCGGC 60
 CGCCGTGGCT ATGTTCTGTG CCGATTTCOG CAAAGAGTTC TACGAGGTGG TCCAGAGCCA 120
 GAGGGTCTCT CTCTCTGTGG CCTCGGACGT GGATGCTCTG TGTGCGTGCA AGATCCTTCA 180
 GGCTTGTTTC CAGGTGACCG ACGTGCAATA TACGCTGGTT CCACTTTCTG GGTGGCAAGA 240
 ACTTGAATCT GCATTTCTTG AGCATAAAGA ACAGTTTCAT TATTTTATTC TCATAAACTG 300
 TGGAGCTAAT GTAGACCTAT TGGATATTCT TCAACCTGAT GAAGACACTA TATCTTTTGT 360
 GTGTGACACC CATAGGCCAG TCAATGTCTG CAATGTATAC AACGATACCC AGATCAAATT 420
 ACTCATTAAG CAAGATGATG ACCTTGAAAT TCCCGCCTAT GAAGACATCT TCAGGGATGA 480
 AGAGGAGGAT GAAGAGCATT CAGGAATGA CAGTGATGGG TCAGAGCCTT CTGAGAAGCG 540
 CACACGGTTA GAAGAGGAGA TAGTGGAGCA AACCATGCGG AGGAGGCGAG GCGAGAGATG 600
 GGAGGCCCGG AGAAGAGACA TCCTCTTTGA CTACGAGCAG TATGAATATC ATGGGACATC 660
 GTGAGCCATG GTGATGTTTG AGCTGGCTTG GATGCTGTCC AAGGACCTGA ATGACATGCT 720
 GTGGTGGGCG ATCGTTGGAC TAACAGACCA GTGGGTGCAA GACAAGATCA CTCAAATGAA 780
 ATACGTGACT GATGTTGGTG TCCTGCAAGC CCAAGTTTCC CGCCACAACC ACCGGAAACG 840
 GATGAGGAG AACCACTCT CCGTGGACTG CACACGGATC TCCTTTGAGT ATGACCTCCG 900
 CCTGGTGCTC TACCAGCACT GGTCCCTCCA TGACAGCCTG TGCAACACCA GCTATACCGC 960
 AGCCAGGCTC AAGCTGTGTG CTGTGCATGG ACAGAAGCGG CTCAGGAGT TCCTTGACGA 1020
 CATGGGTCTT CCCTCAAGAC AGGTGAAGCA GAAGTTCAGG GCCATGGACA TCTCCTTGAA 1080
 GGAGAAATTG CGGGAATGA TTGAAGATC TGCAATAAAT TTTGGGATGA AGGACATGCG 1140
 CGTGCAAGCT TTCAGCATTC ATTTTGGGTT CAAGCACAAG TTTCTGGCCA GCGACGTGGT 1200
 CTTTGCCACC ATGTCTTTGA TGGAGAGCCC CGAGAAGGAT GGCTCAGGGA CAGATCACTT 1260
 CATCCAGGCT CTGGACAGCC TCTCCAGGAG TAACCTGGAC AAGCTGTACC ATGGCCTGGA 1320
 ACTCGCCAGG AAGCAGCTGC GAGCCACCCA CGACACCAAT GCCAGCTGCC TTTGCAACCA 1380
 CCTGTCTATC TCCAGGGGCG CTTTCTGTGA CTGCTCTCTC ATGGAGGGCA CTCAGATGTT 1440
 CATGCTGTTC TCTAGGCCGG CATCCCTAAG CTGCTCTCAG AAACAACCTG TCAAGTCTCT 1500
 TGTGTGTTGG ACAAAGAACC GCGCTGCAA ACTGCTGCCC CTGGTGATGG CTGCCCCCTT 1560

5 GAGCATGGAG CATGGCACAG TGACCGTGGT GGGCATCCCC CCAGAGACCG ACAGCTCGGA 1620
CAGGAGAAC TTTTITGGGA GGGCGTTTGA GAAGGCAGCG GAAAGCACCA GCTCCCGGAT 1680
GCTGCACAAAC CATTITGACC TCTCAGTAAT TGAGCTGAAA GCTGAGGATC GGAGCAAGTT 1740
TCTGCAOGCA CTTATITCCC TCTGTCTCTA GGAATTGAT TCTTCCAGAA TGACCTTCTT 1800
ATTTATGTAA CTGGCTTTCA TTTAGATTGT AAGTTATGGA CATGATTGGA GATGTAGAAG 1860
CCATITTTTA TTAATAAAAA TGCTTATTTT AGGCTCCGTC CCCAAAAAAA AAAAAAAA 1920
AAAAAAA AA

10 Seq ID NO: 253 Protein sequence:
Protein Accession #: NP_003495.1

1 11 21 31 41 51
MFVSDFRKEF YEYVQSORVL LFVADVDAL CACKILQALF QCDHVQYTLV FVSGWQLEET 60
15 AFLEHKEQFH YFILINCGAN VDLIDLQPD EDTIFFVCDT HRPVNVVNVY NDTQIKLLIK 120
QDDLEVPAY EDIFRDEED EESHGNDSDG SEPSEKTRRL EEEIVBQTHR RRQRREWEAR 180
RRDILFDYEQ YEHGTSSAM VMFELAWMLS KDLDMLWNA IVGLTDQWVY DKITQMKYVT 240
DVGVLQRHVS RHNRHNEDEE NTLSDVCTRI SFEDYDLRLVL YQHWLSLHSL CNTSYTAARF 300
20 KLMSVHGQKR LQEFLLADML PLKQVKQKQK AMDISLKENL REMIEESANK FGMKDMRVQT 360
PSIHFGFKH FLASDVVPAT MSLMESPEKD GSGTDHFIQA LDSLSRSNLD KLYHGLELAK 420
KQLRATQQTIT ASCLCTNLVI SQGFPLYCSL MEGTPDVMLF SRPASLSLLS KHLKSPVCS 480
TKNRRCKLLP LVMAAPLSME HGTVTVVGIP PETDSSDRKN FFGRAFEKAA ESTSSRMLHN 540
HFDLSVIELK AEDRSKFLDA LISLLS

25 Seq ID NO: 254 DNA sequence
Nucleic Acid Accession #: NM_022337
Coding sequence: 48..683

30 1 11 21 31 41 51
GGCTGCGCTT CCTGTGCTAG GCACGCGCAG TCTGGCCGGC CGCCAGGATG CAGGCCCGCG 60
ACAAGGAGCA CCGTGTACAG TTGCTGGTGA TTGGCGACCT GGGCGTGGGG AAGACCAGTA 120
35 TCATCAAGCG CTACGTGCAC CAGAACTTCT CCTCGCACTA CCGGGCCACA ATCGGCGTGG 180
ACTTCGCGCT CAAGGTGCTC CACTGGGACC CGGAGACTGT GGTGGCGCTG CAGCTCTGGG 240
ATATGCGAGG TCAGAAAGA TTTGGAAACA TGACGAGGGT CTATTACCGA GAAGCTATGG 300
GTGCATTAT TGCTTCGAT GTCAACAGGC CAGCCACATT TGAAGCAGTG GCAAAGTGA 360
AAAATGATT GGAATCCAG TTAAGTCTCC CTAATGGCAA ACCGTTTCA GTGCTTTGT 420
40 TGGCCAAACA ATGTGACCA GGGAGGATG TGCTCATGAA CAATGGCCTC AAGATGGACC 480
AGTTCTGCAA GGAGCACGGT TTCGTAGGAT GGTTTGAAAC ATCAGCAAG GAAATATAA 540
ACATTGATGA AGCTCCAGA TGCTGTGTA AACACATACT TGCAATGAG TGTGACCTAA 600
TGAGGCTAT TGAGCCGGAC GTCGTGAAG CCCATCTCAC ATCAACCAAG GTTGCCAGCT 660
GCTCTGGCTG TGCCAAATCC TAGTAGGCAC CTTTGTGGT GTCTGTAGG AATGACCTCA 720
45 TTGTTCCACA AATGTGGCT CTATTTTAC CATTTTGGGT AAACGTGAG ATAGATATAC 780
CACATGTGGC AAGCCAAAGA TCTATGCCTC TGTTTTTC ATGAGAGAGA AATAGCAAT 840
GTTCTTTTCA TGCTTCTC ACCATCATCA CAGTGTTC AAACCTTTGA AATATTATG 900
TCTGTACAA ACTCTGTCA TGTAGCTGAC CAAATCTCTG CAGGGCCACA GTCCGCACTG 960
TTATTGCTT CTATTATCA GCAAAGGCT CAAGTCTTAA AATAAAGGG GAGAGAACA 1020
50 AACTAGCTGT CAAGTCAAG ACTGGCTTTC ACCTTGCCCT GGTGTCTTT TCCAGATTTC 1080
AATATATCT CTGATGGCT GACAGGCTTA TTAAGTAGAT GTGATATTT CTTCACAGAT 1140
GACCTCCATT CTGGGAGAC CTAAGAGTTG CCTCTGAGT AGCTCTTGG AATCGTGAAC 1200
ACAGGTGTGC TATATTGTCC TTGCTTAAC TGTCACTGC CATGGCTGA ATGTGCTT 1260
AACTGAATAT TGTATGAAA GACATGCCCT CATATGTGCC TTTCTGTAG CTCTCTTGA 1320
55 CTCAAGCTGT GGGGCTCCTC TATACATGCT ATACATGTAA TATATATTAT ATATATTTT 1380
GCAAGTGAAAC AATAAACAT TAAAGATAA AA

Seq ID NO: 255 Protein sequence:
Protein Accession #: NP_071732

60 1 11 21 31 41 51
MQAPHKEHLY KLLVIGDLGV GKTSIIKRYV HQNFSSHYRA TIGVDFALKV LHWDPETVVR 60
65 LQLWDIAGQE RFGNMTRVYV REAMGAFIVF DVTRPATFEA VAKWKNDLDS KLSLFNGKPV 120
SVVLLANKCD QGKDVLMNNG LKMDQFCKEH GFVGMFETSA KENINIDEAS RCLVKHILAN 180
ECDLMSIEP DVVKPHLTST KVASCSCAK S

70 Seq ID NO: 256 DNA sequence
Nucleic Acid Accession #: NM_016321
Coding sequence: 25..1464

75 1 11 21 31 41 51
GGAACGCGCC GCTGCCAGCC CGGCCAGGCA CCCCTGCAGC ATGGCCTGGA ACACCAACCT 60
CGCTGGCGG CTGCCGCTCA CCTGCCTGCT CCTGCAGGTG ATTATGGTGA TTCTCTTCGG 120
GGTGTTCGTG CGCTACGACT TCGAGGCGGA CGCCCACTGG TGCTCAGAGA GGACGCACAA 180
GAACCTGAGC GACATGGAGA ACGAATTCTA CTATCGCTAC CCAAGCTTCC AGGACGTGCA 240
CGTATGGTTC TTGCTGGGCT TCGGCTTCTT CATGACTTTC CTGCAGGCTG ACGGCTTCAG 300
CGCGTGGGCT TCAACTTCC TGTGGGAGC CTTGGGATC CAGTGGGCGC TGCTCATGCA 360
80 GGGCTGGTTC CACTTCTTAC AAGACCGCTA CATCGTCTGT GCGTGGGAGA ACCTCATCAA 420
CGCTGACTTC TGGCTGGCCT CTGCTGCGT GGCCTTTGGG GCAGTTCGTG GTAAAGTCAG 480
CCCCATTCAG CTGCTCATCA TGACTTCTT CCAAGTGACC CTCTTCGCTG TGAATGAGTT 540
CATCTCTCTT AACCTGCTAA AGGTGAAGGA TGCAGGAGGC TCCATGACCA TCCACACATT 600
TGGCGCTTAC TTTGGGCTCA CAGTGACCGG GATCCTCTAC CGACGCAACC TAGAGCAGAG 660
85 CAAGAGAGA CAGAACTCTG TGTACAGTC GGACCTCTTT GCCATGATTG GCAACCTCTT 720
CCTGTGGATG TACTGGCCCA GCTTCAACTC AGCCATATCC TACCATGGGG ACAGCCAGCA 780
COGAGCGGCC ATCAACACCT ACTGCTCCTT GGCAGCCTGC GTGCTTACCT CGGTGGCAAT 840

ATCCAGTGCC CTGCACAAGA AGGGCAAGCT GGACATGGTG CACATCCAGA ATGCCAGGCT 900
 CGCAGGAGGG GTGGCCGTGG GTACCCGTGC TGAGATGATG CTCATGCCTT ACGGTGCCCT 960
 CATCATCGGC TTGCTCTGGG GCATCATCTC CACCCTGGGT TTTGTATACC TGACCCCAT 1020
 5 CTTGGAGTCC CGGCTGCACA TCCAGGACAC ATGTGGCATT AACAACTGTC ATGGCATTCC 1080
 TGGCATCATA GGGGGCATCG TGGGTGCTGT GACAGCGGCC TCCGCCAGCC TTGAAGTCTA 1140
 TGGAAAAGAA GGGCTTGTCC ATTCTTTTGA CTTTCAAGGT TTCAACGGGG ACTGGACCGC 1200
 AAGAACAAGG GGAAGATTCC AGATTTATGG TCTCTTGGT ACCCTGGCCA TGGCCCTGAT 1260
 GGGTGGCATC ATTGTGGGCG TCATTTTGTG ATTACCATTC TGGGGACAAC CTTTCAGATGA 1320
 10 GAAGCTGCTT GAGGATGCGG TCTACTGGGA GATGCTGAA GGAACAGCA CTGTCTACAT 1380
 CCTGAGGAC CCCACCTTCA AGCCCTCAGG ACCCTCAGTA CCTCAGTAC CCATGGTGTG 1440
 CCCACTACCC ATGGCTTCTT CGGTACCCTT GGTACCCTAG GCTCCAGGG CAGGTGAGGA 1500
 GCAGGCTCCA CAGACTSTCC TGGGGCCAG AGGAGCTGGT GCTGACCTAG CTAGGGATGC 1560
 AAGAGTGAGC AAGCAGCACC CCCACCTGCT GGCTTGGCCT CAAGGTGCCT CCACCCCTGC 1620
 CCTCCCTTTC ATCCAGGGG GTCTGMCCTGA GAATGGAGAA GGAGAAGCTA CAAAGTGGGC 1680
 15 ATCCAAAGCG GGTCTGGCT GCAGAAGTTC TGCCTCTGCC TGGGGTCTTG GCCACATTGG 1740
 AGAAAAACAG GCTCAAAGTG GGGCTGGGAC CTGGTGGGTG AACCTGAGCT CTCCAGGAG 1800
 ACAACTTACG TGCCAGTACC CACCTATGAG GCTCTTCTAC CCGGTGCGCT CACCTGGGCC 1860
 AGCATCTCTT ATGCTCCCTG GGTCCCCAG ACCTCTCTGT GTTGTGTGCG TGGCAGCCTC 1920
 20 CAGGAATAAA CATTCTTGTT GTCCTTTGTA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 257 Protein sequence:
 Protein Accession #: NP_057405

1 11 21 31 41 51
 MAWNTNLNRW LPLTCLLLQV IMVILFGVFP RYDFEADAHW WSETRHNLN DMENEFYRY 60
 PPSODVHVWV FVGFGLMTF LQRYGFSAVG FNFLAAGFI QWALLMQGF HFLQDRYIV 120
 GVENLINADF CVASVCVAPG AVLKGVSPIQ LLIMTFQVLT LEAVNEFILL NLLKVKDAGG 180
 30 SMTIHTPGAY FGLTVTRILY RRLNQSKER QNSVYQSDLP AMIGTLPLWM YWPSFNSAIS 240
 YHGDSDHRAA INTYCSLAAC VLTSAISSA LHKKGKIDMV HIGNATLAGG VAVGTAREMM 300
 LMPYGLIIG FVCGIISTLG PVYLTFFLES RLHIQDTCGI NNLEHIGPIL GGIIVGAVTAA 360
 SASLEFYVGE GLVHSPDFQG FNGDWTARTQ GKPLIYGLLV TLMALMGGI IVGLILRLPF 420
 WQPSDENECF EDAVYWMPE GNSTVYIPED PTFKPSGSPV PSVPMVSLP MASSVFLVP

Seq ID NO: 258 DNA sequence
 Nucleic Acid Accession #: NM_002358.2
 Coding sequence: 75..692

1 11 21 31 41 51
 GGGAAAGTCT GTTGGAGCGG CTGTGGTTGC TGTCCGCGGA GTGGAAGCGC GTGCTTTTGT 60
 TTGTGTCCCT GGCCATGGCG CTGCAGCTCT CCGCGGAGCA GGAATCACC CTGCGCGGGA 120
 45 GGGCCGAAAT CGTGGCCGAG TTCTTCTCAT TCGGCATCAA CAGCATTTTA TATCAGCGTG 180
 GCATATATCC ATCTGAAACC TTTACTCGAG TGCAGAAATA CGGACTCACC TTGCTTGTA 240
 CTACTGATCT TGAGCTCATA AAATACCTAA ATAATGTGGT GGAACCACTG AAAGATTGGT 300
 TATACAAGTG TTCAGTTCAG AAATCGTTG TAGTTATCTC AAATATTGAA AGTGGTGAGG 360
 TCTTGGAAGG ATGGCAGTTT GATATTGAGT GTGACAAGAC TGCAAAAGAT GACAGTGCAC 420
 50 CCAGAGAAAA GTCTCAGAAA GCTATCCAGG ATGAAATCCG TTCAGTGATC AGACAGATCA 480
 CAGCTACGGT GACATTTCG CCACTGTTGG AAGTTTCTTG TTCATTGAT CTGCTGATT 540
 ATACAGACAA AGATTGGTT GTACCTGAAA AATGGGAAGA GTCCGGACCA CAGTTTATTA 600
 CCAATTCGA GGAAGTCCGC CTTCGTTTCT TACTACTAC AATCCACAAA GTAAATAGCA 660
 TGGTGGCCTA CAAATTCCT GTCAATGACT GAGGATGACA TGAGGAAAT AATGTAATTG 720
 55 TAAATTTGAA ATGTGGTTTT CTGAAATCA GGTCACTAT AGTTGATATG TTTTATTTCA 780
 TTGGTTAATT TTTACATGGA GAAACCAAAA ATGATACTTA CTGAACTGTG TGTAAATTGT 840
 CCTTATTTT TTTGGTACCT ATTTGACTTA CCATGGAGTT AACATCATGA ATTTATTGCA 900
 CATTTGTTCA AAGGAACAG GAGGTTTTTT TGTCAACATT GTGATGTATA TTCTTTGAA 960
 60 GATAGTAATC GTAGATGGA AACTTGTGC TATAAAGCTA GATGCTTCC TAAATCAGAT 1020
 GTTTTGTGCA AGTAGTTTGA CTCAGTATAG GTAGGAGAT ATTTAAGTAT AAAATACAA 1080
 AAAGGAAGTC TAAATATCA GAATCTTTGT TAAGGTCCTG AAAGTAACCT ATAATCTATA 1140
 AACATGAAA TATGTCTGTA TAGCTCCTTT TGACCTTCAT TTCATGTATA GTTTTCCCTA 1200
 TTGAATCAGT TTCCAATTAT TTGACTTTAA TTTATGTAAC TTGAACCTAT GAAGCAATGG 1260
 ATATTTGTAC TGTTTAATGT TCTGTGATAC AGAACTCTTA AAAATGTTTT TTCATGTGTT 1320
 65 TTATAAATC AAGTTTTAAG TGAAAGTGAG GAAATAAAGT TAAGTTTGT TTAATAAAAA 1380
 AAAAAAAAAA

Seq ID NO: 259 Protein sequence:
 Protein Accession #: NP_002349.1

1 11 21 31 41 51
 MALQLSREQG ITLRGSABIV AEFFSFGINS ILYQRGIYPS ETPTRVQKYG LTLVTTDL 60
 75 LIKYLANNVE QLKDWLYKCS VQKLVVVISN IESGEVLERN QPDIECDKTA KDDSAPEKRS 120
 QKAIQDEIRS VIRQITATVT FLPLLEVSCS PDLLIYTDKD LVPPEKWEES GPQFITNSEE 180
 VRLRSFTTTI HKVNSMVAYK IPVND

Seq ID NO: 260 DNA sequence
 Nucleic Acid Accession #: NM_001211
 Coding sequence: 43..3195

1 11 21 31 41 51
 85 AAAGGCTGCG AGCAGGACGA GGACCTGAGC CAGGAATGCA GGATGGCGGC GGTGAAGAG 60
 GAAGGGGGTG CTCTGAGTGA AGCCATGTCC CTGGAGGAG ATGAATGGGA ACTGAGTAA 120
 GAAATGTAC AACCTTTAAG GCAAGGGCGG ATCATGTCCA CGCTTCAGGG AGCACTGGCA 180
 CAAGAATCTG CCTGTAACAA TACTCTTCAG CAGCAGAAAC GGCATTGTA ATATGAAATT 240

	CGATTTTACA	CTGGAATGA	CCCTCTGGAT	GTTTGGGATA	GGTATATCAG	CTGGACAGAG	300
	CAGAACTATC	CTCAAGGTGG	GAAGAGAGT	AATATGTCAA	CGTTATTAGA	AAGAGCTGTA	360
	GAAGCACTAC	AAGGAGAAAA	ACGATATTAT	AGTGATCCTC	GATTTCTCAA	TCTCTGGCTT	420
5	AAATTAGGGC	GTTTATGCAA	TGAGCCTTGG	GATATGTACA	GTTACTTGCA	CAACCAAGGG	480
	ATTGGTGT	CACTTGCTCA	GTTCTATATC	TCATGGGCGAG	AAGAATATGA	AGCTAGAGAA	540
	AACTTTAGGA	AAGCAGATGC	GATATTTTAC	GAAGGGGATT	AACAGAAAGC	TGAACCACTA	600
	GAAGAGCTAC	AGTCCCAACA	CCGACAATTC	CAAGCTCGAG	TGCTCGGCA	AACCTCTGTT	660
	GCACCTTGAG	AAGAAGAAGA	GGAGGAAGTT	TTTGAGTCTT	CTGTACCACA	ACGAAGCACA	720
10	CTAGCTGAAC	TAAAGAGCAA	AGGGAAGAA	ACAGCAAGAG	CTCCAATCAT	CGGTGTAGGA	780
	GGTGCTCTCA	AGGCTCCAAG	CCAGAACAGA	GGACTCCAAA	ATCCATTTC	TCAACAGATG	840
	CRAAATAATA	GTAGAATTAC	TGTTTTTGAT	GAATAATGCTG	ATGAGGCTTC	TACAGCAGAG	900
	TTGTCTAAGC	CTACAGTCCA	CCCATGGATA	GCACCCCA	TGCCAGGGC	CAAGAGAAAT	960
	GAGCTGCAAG	CAGGCCCTTG	GAACACAGGC	AGGTCTCTGG	AACACAGGCC	TCGTGGCAAT	1020
15	ACAGCTTCAC	TGATAGCTGT	ACCCGCTGTG	CTTCCAGATT	TCCTCCATA	TGTGGAAGAG	1080
	ACTGCACAAC	AGCCAGTTAT	GACACCATGT	AAAATTGAAC	CTAGTATAAA	CCCATCCTTA	1140
	AGCACCAGAA	AGCCTGGAAA	GGAGAAGGA	GATCCTCTAC	AAAGGGTTCA	GAGCCATCAG	1200
	CAAGCGTCTG	AGGAGAAGAA	AGAGAAGATG	ATGTATTGTA	AGGAGAAGAT	TTATGCAGGA	1260
	GTAGGGGAAT	TCTCCTTGA	AGAAATTCGG	GCTGAAGTTT	TCCGGAAGAA	ATTAAAGAG	1320
20	CAAGGGGAAG	CCGAGCTATT	GACCAAGTGA	GAGAAGAGAG	CAGAAATGCA	GAACAGATT	1380
	GAAGAGATGG	GAGTTTATGC	AAAAGAAATC	CAAACTACTC	AGCAAGAAAG	AACAGGTGAT	1440
	CAGCAAGAG	AGACGATGCC	TACAAAGGAG	ACAACTAAAC	TGCAAAATGC	TCCGAGTCT	1500
	CAGAAATATC	CAGGAATGAC	TCTATCCAGT	TCTGTTTGTC	AAGTAACTG	TTGTGCCAGA	1560
	GAACCTTCAC	TTCGGAGAA	CATTGGCAG	GAACAACCTC	ATTCTAAAGG	TCCAGTGTAT	1620
25	CCTTTCTCCA	TTTTTGATGA	GTTTCTTCTT	TCAGAAAAGA	AGAATAAAG	TCCTCTGCA	1680
	GATCCCCCAC	GAGTTTATGC	TCAACGAAGA	CCCTTCAG	TTCTCAAAAC	CTCAGAAAGC	1740
	ATCACTCAA	ATGAAGATGT	GTCTCCAGAT	GTTTGATGAT	AATTACAGG	AATTGAACCC	1800
	TTGAGCGAGG	ATGCCATTAT	CACAGGCTTC	AGAAATGTAA	CAATTTGTCC	TAACCCAGAA	1860
	GACACTTGTG	ACTTTGCCAG	AGCAGCTCGT	TTTGTATCCA	CTCCTTTCA	TGAGATAATG	1920
30	TCCTTGAAGG	ATCTCCCTTC	TGATCCTGAG	AGACTGTTAC	CGGAAGAAGA	TCTAGATGTA	1980
	AAGACCTCTG	AGGACCAACA	GACAGCTTGT	GGCACTATCT	ACAGTCAGAC	TCTCAGCATC	2040
	AAGAAGCTGA	GCCCAATTAT	TGAAGACAGT	CGTGAAGCCA	CACACTCCTC	TGGCTTCTCT	2100
	GGTTCTTCTG	CCTCGGTTGC	AAGCACCTCC	TCCATCAAAT	GTCTTCAAAT	TCCTGAGAAA	2160
	CTAGAACCTA	CTAATGAGAC	TTCAGAAAAC	CCTACTCAGT	CACCATGGTG	TTCACAGTAT	2220
35	CGCAGACAGC	TACTGAAGTC	CCTACCAGAG	TTAAGTGCCT	CTGCAGAGTT	GTGTATAGAA	2280
	GACAGACCAA	TGCTTAAGTT	GGAAATTGAG	AGGAAATTG	AATTAGGTAA	TGAGGATTAC	2340
	TGCATTAAAC	GAGAATACTT	AATATGTGAA	GATTACAAGT	TATCTGGGT	GGGCCCAAGA	2400
	AACCTGTCAG	AATTACAGT	AATAAAGGTA	TCTTCTCAAC	CTGTCCCATG	GGACTTTTAT	2460
	ATCAACCTCA	AGTTAAAGGA	ACGTTTAAAT	GAAGATTG	ATCATTTTGT	CAGCTGTTAT	2520
40	CAATATCAAG	ATGGCTGTAT	TGTTTGGCAC	CAATATATAA	ACTGCTTCAC	CCCTCAGGAT	2580
	CTTCTCCAAC	ACAGTAGAATA	TATTACCCAT	GAATAACAG	TGTTGATTAT	TTATAACCTT	2640
	TTGACAATAG	TGGAGATGCT	ACACAAAGCA	GAATAGTCC	ATGGTGACTT	GAGTCCAAGG	2700
	TGTTCTGATT	TCAGAAACAG	ATCCACGAT	CCCTATGATT	GTACAAAGAA	CAATCAAGCT	2760
	TTGAAGATAG	TGGACTTTTC	CTACAGTGT	GACCTTAGGG	TGCAGCTGGA	TGTTTTTACC	2820
45	CTCAGCGGCT	TTCGCACTGT	ACAGATCCTG	GAAGGACAAA	AGATCCTGGC	TAACCTGTTCT	2880
	TCTCCCTACC	AGTAGAGACTT	GTTTGGTATA	GCAGATTAG	CACATTACT	ATTGTTCAAG	2940
	GAACACCTAC	AGGTCTTCTG	GGATGGGTCC	TTCTGGAAAC	TTAGCCAAA	TATTTCTGAG	3000
	CTAAAGATG	GTGAATTGTG	GAATAAATTC	TTTGTGCGGA	TTCTGAATGC	CAATGATGAG	3060
	GCCACAGTGT	CTGTTCTTGG	GGAGCTTGCA	GCAGAAATGA	ATGGGGTTTT	TGACACTACA	3120
50	TTCCAAAGTC	ACCTGAACAA	AGCCTTATGG	AAGTAGGGA	AGTTAACTAG	TCCTGGGGCT	3180
	TGCTCTTTT	AGTGTAGCTAG	GCAATCAAGT	CTCACAGATT	GCTGCCCTCAG	AGCAATGGTT	3240
	GTATTGTGGA	ACACTGAAC	TGTATGTGCT	GTAATTTAAT	TTAGGACACA	TTTAGATGCA	3300
	CTACCATGTC	TGTTCTACTT	TTTGTGACAG	GTATATTTG	ACGTCACTGA	TATTTTTTAT	3360
	ACAGTGATAT	ACTTACTCAT	GGCCTTGCT	AACCTTTTGT	AAGAACTATT	TTATTCTAAA	3420
55	CAGACTCATT	ACAAATGTTT	ACCTTGTTAT	TTAACCCATT	TGCTCTACT	TTTCCCTGTA	3480
	CTTTTCCCAT	TGTGAATTG	TAAAAATGTC	TCTTATGATC	ACCATGTATT	TTGTAAATAA	3540
	TAAATAGTA	TCTGTTAAAA	AAAAAATAA	AAAAAATAA	AAA		

Seq ID NO: 261 Protein sequence:
Protein Accession #: NP_001202

60	1	11	21	31	41	51	
	MAAVKKEGGA	LSEAMSLBEG	EWELSKENVQ	PLRQGRIMST	LQGALAQESA	CNNTLQQQKR	60
65	AFEYBIRFYT	GNDPLDVWDR	YISWTEQNYF	QGGKESNMST	LLERAVEALQ	GEKRYSDPR	120
	FLNLNLKLR	LONEPLDMYS	YLNQIGIGVS	LAQFYISWAB	EYEARENFRK	ADAIFQEGIQ	180
	QKAEPLERLQ	SQHRFQARV	SRQTLALEK	EEEEVEFESS	VPQRSTLAEI	KSRGKKTARA	240
	PIIRVGGLK	APSQNRGLQN	PPQQMQNNS	RITVFDENAD	EASTAELSKP	TVQFWIAPPM	300
	PRAKENELQA	GPWNTGRSLE	HRPRGNTASL	IAPFAVLPSF	TPYVEETAQQ	PVMTPKIEP	360
70	SINHLSTRK	PGKEEGDPLQ	RVQSHQQAEE	EKKKEKMYCK	EKIYAGVGFE	SFEIRAEVF	420
	RKKLKEQREA	ELLTSAEKRA	EMQKQIEEME	KKLKEIQTTQ	QERTGDQOEE	TMPTKETTKL	480
	QIASESQIRP	GMLTSSSVCC	VNCCARETSL	AEINIWEQPH	SKGPSVFFSI	FDEPLLSEKK	540
	NKSPADPPR	VLAQRRLP	LKTSSESITN	EDVSPDVCE	FTGIEPLSED	AIITGFRNVT	600
	ICPNPEDTC	PARAARFVST	PFHEIMSLKD	LPSDPERLLP	EEDLDVKTSE	DQQTACGTIY	660
75	SQLSLIKKLS	PILEDSEAT	HSSGFGSSGA	SVASTSSIKC	LQIPEKLELT	NETSENPTQS	720
	PWCSQYRRQL	LKSLPELSAS	AELCIEDREM	PKLEIEKEIE	LGNEDYCIKR	EYLICEDYKL	780
	PWVAPRNSAE	LTIVKVSQF	VPWDFYINLK	LKERLNEFD	HFCSCYQYQD	GCIVWHQYIN	840
	CFTLQDLQHH	SEYITHEITV	LIINYLLTIV	EMLHKAEIVH	GDLSRCLIL	RNRHDPYDC	900
	NKNNQALXIV	DFYSVDLRV	QLDVFTLSGF	RTVQILEGQK	ILANCSSPYQ	VDLFGIADLA	960
80	HLLEPKHEHQ	VFDGFSFWKL	SONISELKDG	ELWNKFFVRI	LNANDEATVS	VLGELAAEMN	1020
	GVFDTPPQSH	LKALWKVGK	LTPGALLFQ				

Seq ID NO: 262 DNA sequence
Nucleic Acid Accession #: NM_003784
Coding sequence: 365..1507

85	1	11	21	31	41	51
----	---	----	----	----	----	----

	GTCTACTTAT	CAATAAGCAG	CTGCTGTGTC	AGAGTGCAGG	CTGCACCTTT	GGACAGCCTT	60
	TAAAACTGAA	TTCTCAGAA	TTTAGAACAA	ATTTTGTGCT	AGAAATGCTG	ACTTTGGTTC	120
5	ATTAGTAGT	GGTAAACACAG	GCTCCCTTCG	AAGCTCTCCT	TCATCACCTT	CCTAAGTGCA	180
	TGTACAGGGA	AGCTCTCCTT	CATCACCTTC	CTAAGTGCAT	GGGGGAAAT	ACCTAGGGCT	240
	CAACAGTCTT	GAGAAGTGTG	GAAACATTTT	CTTTGTGAGT	GAGAACAGAT	CACCTAGAGA	300
	AAGGAAACCA	GATTCCCATC	ACTGCTTCTG	GGTATCAGAT	GCTAGCGCTG	CACCTCCATT	360
	TGCAATGGCC	TCCTCTGCTG	CAGCAAAATG	AGAGTTTTGC	TTCAACCTGT	TCAGAGAGAT	420
10	GGATGACAA	CAAGGAAATG	GAAATGTGTT	CTTTTCTCTC	CTGAGCCTCT	TCGCTGCCCT	480
	GGCCCTGGTC	CGCTTGGGGG	CTCAAGATGA	CTCCCTCTCT	CAGATTGATA	AGTTGCTTCA	540
	TGTTAACT	GCTCAGGAT	ATGGAAGTCT	TTCTAATAGT	CAGTCAGGGC	TCCAGTCTCA	600
	ACTGAAACCA	GTTTCTTCTG	ATATAAATGC	ATCCCAACAG	GATTATGATC	TCAGCATTGT	660
	GAATGGGCTT	TTTGTGAAA	AAGTGTATGG	CTTTCATAAG	GACTACATTG	AGTGTGCGGA	720
15	AAAATTATAC	GATGCCAAAG	TGGAGCGAGT	TGACTTTACG	AATCATTAG	AAGACACTAG	780
	ACGTAATATT	AATAAGTGGG	TTGAAAATGA	AACACATGCG	AAAATCAAGA	ACGTGATTGG	840
	TGAAGGTGGC	ATAAGCTCAT	CTGCTGTAAT	GGTGTGCTG	AATGCTGTGT	ACTTCAAAGG	900
	CAAGTGGCAA	TCAGCCTTCA	CCAAGAGCGA	AACCATAAAT	TGCCATTTC	AATCTCCCAA	960
	GTGCTCTGGG	AAGGCAGTCG	CCATGATGCA	TCAGGAACGG	AAGTTCAATT	TGTCTGTTAT	1020
20	TGAGGACCCA	TCATGAAGA	TTCTTGAGCT	CAGATACAAT	GGTGGCATA	ACATGTACGT	1080
	TCTGCTGCCT	GAGAATGACC	TCTCTGAAAT	TGAAAAACAA	CTGACCTTTC	AGAATCTAAT	1140
	GGATGGACC	AATCCAAAGC	GAATGACCTC	TAAGTATGTT	GAGGTATTTT	TTCTCAGTT	1200
	CAAGATAGAG	AAGAATTTATG	AAATGAAACA	ATATTGAGA	GCCCTAGGGC	TGAAAGATAT	1260
	CTTTGATGAA	TCCAAAGCAG	ATCTCTCTGG	GATTGCTTGG	GGGGGTCTGC	TGTATATATC	1320
25	AAGGATGATG	CACAAATCTT	ACATAGAGGT	CACAGAGGAG	GGCACCGAGG	CTACTGCTGC	1380
	CACAGGAAGT	AATATTGTAG	AAAAGCAACT	CCCTCAGTCC	ACGCTGTTTA	GAGCTGACCA	1440
	CCCATTCTTA	TTTGTATCA	GGAAGGATGA	CATCATCTTA	TTCACTGGCA	AAGTTTCTTG	1500
	CCCTTGAAA	TCCAATTGTT	TTCTGTTATA	GCAGTCCCCA	CAACATCAAA	GRACCAACCAC	1560
	AAGTCAATAG	ATTTGRTGTT	AATGGGAAA	ATGTGGTGT	TCCTTTGAGT	TTATTTCTTC	1620
30	CTAACATTGG	TCAGCAGATG	ACACTGGTGA	CTTGACCTTT	CCTAGACACC	TGCTTGATTG	1680
	TCCTGATGCC	TGCTCTTAGC	ATTCTACCC	CATGTGTCTC	ACCCATTCTT	AATTTTATTG	1740
	TCTTTCTTCC	CAGCTCATT	TCTATCATT	TCCCCCATGA	CCGCTCTGGA	AATTATGGAG	1800
	RGTCCTCAAC	TGGTAAGGAG	AACGTAGAAG	TAGCCCTAGG	GATCCTTTT	GAAACTCTAC	1860
	AGTTATCGCA	GATATTCTAG	CTTCATTGTA	AGCAATCTAG	GAAATAAGCC	CTGCTGCTTT	1920
35	CTAGAAATAA	GTGTGAAGGA	TAAATTTTCT	TTGTTGACCT	ATGAAGATT	TAGAGTTTAC	1980
	CTTCATATGT	TTGATTTTAA	ATCAGTGTAT	AATCTAGATG	GTAAAAAATG	TGAAATTGGG	2040
	ATTAGGAGCC	TACCAAAATA	TTTCATTAAT	GCTTTCAATT	GACAAATTTT	GGCCTTTCTT	2100
	TGATAAGACA	ATATGTACAT	GTTTTTTC	ATATTAAAGA	TCTTTTAACT	GTTGGCAGTT	2160
40	GTTATCTACA	GAATCATATT	TCATATGCTG	TGTAGTTTAT	AAGTTTTC	TCTATTATC	2220
	AGAATAAAGA	AATACACAT	ACCTGTAA				

Seq ID NO: 263 Protein sequence:
Protein Accession #: NP_003775

45	1	11	21	31	41	51	
	MASLAANAE	PCPNLFREMD	DNQNGNVFF	SSLSLFAALA	LVRLGAQDD	LSQIDKLLHV	60
	NTASGYGNSS	NSQSGLSQSL	KRVFSDINAS	HKDYDLISVN	GLPAEKVYGF	HKDYIECAEK	120
50	LYDAKVERVD	FTNHLEDTRR	NINQWVENET	HGKIKNVIGE	GGISSAVMV	LVNAVYPKGG	180
	WQSAPFKSET	INCHFSKPKC	SGKAVAMHQ	ERKFNLVIE	DPSMKILELR	YNGGINMYVL	240
	LPENDLSEIE	NKLTQNLME	WTNPRMTSK	YVEVFPQFK	IEKNYEMKQY	LRALGLKDIF	300
	DESKADLSGI	ASGGRLYISR	MMHKSIEVT	EEGTETAAT	GSNIVEKQLP	QSTLFRADHP	360
	FLFVIRKDDI	ILPSGKVSCE					

Seq ID NO: 264 DNA sequence
Nucleic Acid Accession #: AB052906
Coding sequence: 74-814

60	1	11	21	31	41	51	
	AAAACCTTGA	GGTGATTCAT	CTTCCAGGCT	CTCCTTCCAT	CAAGTCTCTC	CTCCCTAGCG	60
	CTCTGGGTCC	TTAATGGCAG	CAGCCGCGCG	TACCAAGATC	CTTCTGTGCC	TCCCGCTTCT	120
	GCTCCTGCTG	TCCGGCTGGT	CCCGGGCTGG	GCGAGCGGAC	CCTCAGTCTC	TTTGCTATGA	180
65	CATCACCGTC	ATCCCTAAGT	TCAGACCTGG	ACCACGGTGG	TGTGCGGTTC	AAGGCCAGGT	240
	GGATGAAAAG	ACTTTTCTTC	ACTATGACTG	TGGCAACAAG	ACAGTCACAC	CTGTCACTCC	300
	CCTGGGAAG	AAACTAAATG	TCACACGGCG	CTGGAAAGCA	CAGAACCCAG	TACTGAGAGA	360
	GGTGGTGGAC	ATACTTACAG	AGCAACTGCG	TGACATTCAG	CTGGAGAATT	ACACACCCAA	420
	GGAAACCCCT	ACCCTGCAGG	CCAGGATGTC	TTGTGAGCAG	AAAGCTGAAG	GACACAGCAG	480
70	TGGATCTTGG	CAGTTCAGTT	TOGATGGGCA	GATCTTCTCT	CTCTTTGACT	CAGAGAAGAG	540
	AATGTGGACA	ACGGTTCATC	CTGGAGCCAG	AAAGATGAAA	GAAAAGTGGG	AGAATGACAA	600
	GGTGTGGGCC	ATGTCTCTCC	ATTACTTCTC	AATGGGAGAC	TGTATAGGAT	GGCTTGAGGA	660
	CTTCTTGATG	GGCATGGACA	GCAACCTGGA	GCCAAGTGCA	GGAGCACCAC	TGCCATGTC	720
	CTCAGGCACA	ACCCAATCTA	GGGCCACAGC	CACCAACCTC	ATCCTTTGCT	GCCTCTCAT	780
75	CATCCTCCCC	TGCTTCATCC	TCCCTGGCAT	CTGAGGAGAG	TCCTTTAGAG	TGACAGGTTA	840
	AAGCTGATAC	CAAAAGGCTC	CTGTGAGCAC	GGTCTTGATC	AAACTCGCCC	TTCTGTCTGG	900
	CCAGCTGCCC	ACGACCTACG	GTGTATGTCC	AGTGGCCTCC	AGCAGATCAT	GATGACATCA	960
	TGGACCCAAT	AGCTCATCTA	CTGCTTGTAT	TCCTTTTGCC	AACAATTTTA	CCAGCAGTTA	1020
	TACCTAACAT	ATTATGCAAT	TTTCTCTTGG	TGCTACCTGA	TGGAATTCCT	GCATCTAAAG	1080
80	TTCTGGCTGA	CTAAACAAGA	TATATCATTT	TCTTTCTTCT	CTTTTGTGTT	GGAAATCAA	1140
	GTACTTCTTT	GAAATGATAT	CTCTTCTTCT	CAAAATGATAT	TGTCAATAAA	ATAATCAAGT	1200
	TAGACTTCAG	ACCTCTGGGG	ATTCTTTCCG	TGTCTGAAA	GAGAAATTTT	AAATATTATA	1260
	ATAAGAAAAA	ATTTATATTA	ATGATTGTTT	CCTTTAGTAA	TTTATTGTTT	TGTACTGATA	1320
85	TTTAAATAAA	GAGTTCTATT	TCCCAAAAAA	AAAAA			

Seq ID NO: 265 Protein sequence:
Protein Accession #: BAB61048.1

1 11 21 31 41 51
MAAAATKIL LCLPLLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGGVDEKT 60
FLHYDCGNKT VTPVSLGKX LNVTTAKAQ NPVLREVVDI LTELRLDIQL ENYTPKEPLT 120
5 LQARMSCEQK AEGHSSGSWQ FSPDQIPLL FDSEKRMWTT VHPGARIMKE KWENDKVVM 180
SPHYFPMGDC IGWLEDFLMG MDSTLEPSAG APLAMSSGTT QLRATATTLI LCCLLIILPC 240
FILPGI

10 Seq ID NO: 266 DNA sequence
Nucleic Acid Accession #: XM_084853.1
Coding sequence: 127-444

1 11 21 31 41 51
15 ATTGATGATA TATTAAACGA AATCAAATTT GGTGAATATG TGGACACTGG AAAGCTAATC 60
GACAAGATCA ACTTACCAGA TTCTCTAAAA GTGTACCTTA ACCCAAGGCC ACCTTTTGGT 120
AACACCATGA GTGGCATCCA CAAGAGCTTT GAGGTGCTCG GTTATACCAA CTCCTAAGGG 180
20 AAAAGAGCCA TTGGAAGAGA GGAATCTCTG AGACTGCTCG TTAATAAGAG TGAGCATATG 240
ACGGAGGAGG AGATGTTGGA TTGCTTTGCT TCACTGTTTG GCCTGAATCC CGAGGGATGG 300
AAATCCGAGC CTGCAACCTG CTCGTCAAAA GGTTTCAAAA TTTGCTTGA AGAAGAACTT 360
CCAGACGAAA TCACTGAGCA AATATTCGCG ACTGAAATTC TTGGCTTAAC CATTTCAGAA 420
GATTCGGGCC AGGATGGTCA GTGAAGTTAC CAGGAATGTT TAAAGCACAA AGGACTTTGG 480
GTGTGTGTGC ATGCACATGT GTGTGTTTTC CATGAGGCAC TGCCTTTTAT GCATTTCCCT 540
25 CCCCCCTCTC ATCTTTAGAA CATTTAGACA TTAAGCAAG TTTCTGGTGA GCAATG

30 Seq ID NO: 267 Protein sequence:
Protein Accession #: XP_084853.1

1 11 21 31 41 51
MSGIHKSEFV LGYTNKSGKK AIRREDPLRL LVTKGEHMTB EEMLDCEASL PGLNPEGWKS 60
EPATCSVKGS EICLEEELPD BITAEIPATE ILGLTISEDG GQDQG

35 Seq ID NO: 268 DNA sequence
Nucleic Acid Accession #: NM_001898
Coding sequence: 57-482

1 11 21 31 41 51
40 GGCTCTCACC CTCTCTCTCT GCAGCTCCAG CTTTGTGCTC TGGCTCTGAG GAGACCATGG 60
CCAGATATCT GAGTACCCTG CTGCTCCTGC TGGCCACCCT AGCTGTGGCC CTGGCCTGGA 120
GCCCAAGAGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCAATGATG 180
45 AGTGGGTACA GCGTGCCTTT CACTTCGCCA TCAGCGAGTA TAACAAGGCC ACCAAAGATG 240
ACTACTACAG ACCTCCGCTG CGGTTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGTGA 300
ATTACTTCTT CGAGCTAGAG GTGGGCGGCA CCATATGTAC CAAGTCCGAG CCCAACTTGG 360
ACACCTGTGC CTTCCTATGA CAGCCAGAAC TGCAGAGAAA ACAGTTGTGC TCTTTCGAGA 420
TCTACGAAGT TCCTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCCT 480
50 AGGGATCTGT GCGAGGCCAT TCGCACCAGC CACCACCCAC TCCACCCCTC TGTAGTGCTC 540
CCACCCCTGG ACTGGTGGCC CCCACCTGCG GGGAGGCGCT CCCATGTGCC TGCGCCAAGA 600
GACAGACAGA GAAGGCTGCA GGAGTCTCTT GTTGCTCAGC AGGGCGCTCT GCCCTCCCTC 660
CTTCTCTCTT GCTTCTAATA GCCCTGGTAC ATGGTACACA CCCCCCACC TCTTGCATTT 720
AACAGTAGC ATGCCC

55 Seq ID NO: 269 Protein sequence:
Protein Accession #: NP_001889.1

1 11 21 31 41 51
60 MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWQORALHF AISEYNKATK 60
DDYYRRPLEV LRARQQTGVG VNYFFDVEVG RTICTKSQPN LDTAPHEQP ELQKKQLCSF 120
BIYEVPWENR RSLVKSRCQE S

65 Seq ID NO: 270 DNA sequence
Nucleic Acid Accession #: XM_093210
Coding sequence: 13-1854

1 11 21 31 41 51
70 ATGGCAAGCG CCGGAATCTC CTCAGCTGCC GTTTCACAAA AGAGGTACCA GGTCCGCACC 60
AAACGAGCAC ACAAGCAGCA CCAGGAGCTG CAGAAGAAGG AGGGGCGAGC GATGGACCAG 120
GGCAGAGGGA ATGGGGAGGG GGCATCTCTC CCCATATCTG AGGTGCGACT CGCGGACGTA 180
75 GAGCGGACTG GGCTTTTCCC GTTGGGCGGT GGCTCAATC AGGACTTCTT GCCCACGTGC 240
GCCTTCAAAA CGGTAAGAGC TGCAACTGAA CGTGTGAGAC ATGGTGAGA TAGGCTGAGA 300
GGCGGCGGGA GAGATGCCCA TGAACCTAAG TACCCGAGCA CGCCCTCCAC TTCTACCACC 360
ACGAGTAACA CCGCCCCCAC GGGACCGCTC TCGAGGTCCC CCAAGCCAG GACCGAAGGA 420
GGAAACGCCC GGGCGCGGCG CAGCAGCGGC GGGCACCGGC CCAATGGCCA CGGAATCAG 480
CACTGGCACT CGCGCTCTCT CACACCGCAG CGGTGCAGTG TGGCGAGCGG AGCCTCCCGG 540
80 GCGGAGGACC CAGCTAGGCC GTCAACCCCGG TTGCTCCAC GGAAGGGGCG ACCAGGCAAA 600
CTGCCCAAGG CCGCGAGCCC AGGCTCCCTG GCGGAGGCGT CCGCTGGTCC CGCCAGATC 660
ATGGCCGCCA CCAGGCTCCC GAGCATGGC TTCTGTCCG GGAACGGCCC GGCGTCTGTG 720
CTGTCCAGCT AG

85 Seq ID NO: 271 Protein sequence:
Protein Accession #: XP_093210

1 11 21 31 41 51

5 MLRHGEQKRR RARKKWDFLP TCAFKTVRAA TERVRHGADR LRGGRDAHE LKYPDTPSTS 60
 TTTSTNAPTQ PLRSRSPKPRP QGGTPRRRPA AAGTRANGHG TQHWQSALLT PQACSVADGA 120
 SRAEDPARPS PRLLPREGAP GKLPKAPSPG SLAEASAGLL AHVRLQNADA QRVSIQALP 180
 PNSSVGRKEE RPAGAGQRRR PAMPATELST GSRPSSHRRR AVWPTEPPGP RTQLEPSFRL 240
 LPREGAPGKL PKAPSPGSLA EASAGPAQIM AATRLPSRGP LSGNGPASWL SS

Seq ID NO: 272 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..732

10 1 11 21 31 41 51
 15 GGATACTGTG TCACTCAAAG TAATGGGAGG GAGAGAGAAC AGGGAGGGTA GGGATGCTTT 60
 TGAAAAAGCT TTTTTCCTCA CTTTTAACTT GCTTTAGCGT TAAGAGTACT TACCAGCTAA 120
 TAATGTGGAG GAAATTAITC TTTCTCATGT GAGATTACAG AATATATCTA TTCTCTTGA 180
 ATACCCACTT GAAGCCTCTG TAGAAATGTC TCGTCCCTCG GTTGTATTTC TAAACCTTAC 240
 ATGATTTTGT CTGTGTTCTG CAGTGAGAAA TTACATCCAT AGCAAAGACA AAAGTCTTTT 300
 20 TAAATATATT TTATTTATCT TTATATAGT TCTTACAATT TCTAAAAAT TAACACTCAT 360
 TTAGTATCAC AATTATATGG AGAGGGTTT TGTATTTT AAGCATATGT GGCTTATATA 420
 AAAATTGCAG AAGTCATAGG ACTGTCATGT ATTGCAGCTC TGAGAACCAA TGCTGAAAC 480
 TTAAGCC

Seq ID NO: 273 Protein sequence:

Protein Accession #: Eos sequence

30 1 11 21 31 41 51
 MGGRENREGR DAPEKAPFFT FNLL

Seq ID NO: 274 DNA sequence

Nucleic Acid Accession #: NM_003976.2

Coding sequence: 299-961

35 1 11 21 31 41 51
 40 CTCTGAGCTT CTCTGAGCCT TGTGCTGCTCA TCTGGAAAAA GGGGATTAAA CCATTACCT 60
 CATGGAGTTG TGAAGAATA GCTGCAAGC ACCTAACACA TAGTAAGGTT CCCAGTGCAG 120
 CTACTTCTGC TGGGTTGAGT CTAGCTGTGT AGGCCCTTGT TTCCTCACCT GGAGAACTG 180
 GGGTGGCAGG CCGGTCCCCC ACAAAGATA ACTCATCTCT TAATTGCAA GCTGCCTCAA 240
 CAGGAGGGTG GGGGAACAGC TCAACAATGG CTGATGGGCG CTCTGTGTGT TGATAGAGAT 300
 GGAACCTGGA CTGTGAGGCC TCTCCACGCT GTCCCACTGC CCCTGGCCTA GCGCGCAGCC 360
 45 TGCCCTGTGG CCCACCTGCG CCGCTCTGCG TCTGCTGAGC AGCGTCGCAG AGGCCTCCCT 420
 GGGCTCCGCG CCGCGCAGCG CTGCCCCCGG CGAAGGCCCG CGCGCTGTCC TGGCGTCCCC 480
 CGCCGCGCAC CTGCGGGGGG GACGCACGCG CCGCTGTGTC AGTGAAGAG CCGCGCGGCC 540
 GCGCGCGCAG CCTTCTCGGC CCGCGGCCCG GCGCCTGCA CCCCATCTG CTCTTCCCGG 600
 CGGCGCGCGC GCGCGCGCGG CTGGGGGCCG GGGCAGCGCG GCTCGGCGAG CCGGGCGCGG 660
 50 GGGCTGCCCG CTGCGCTGCG AGCTGGTGCC GGTGCGCGCG CTGCGCTGCG GCCACCGCTC 720
 CGACGAGCTG GTGCGTTTCC GCTTCTGCGAG CGGCTCCTGC CGCGCGCGCG GCTCTCCACA 780
 CGACCTCAGC CTGGCCAGCC TACTGGGCGC CGGCGCCCTG CGACCGCCCC CGGGCTCCCG 840
 GCGCGTGCAG CAGCCCTGCT GCGGACCCAC GCGCTACGAA GCGGTCTCCT TCATGACGT 900
 CAACAGCACC TGGAGAACCG TGGACCGCCT CTCGCCACCG GCCTGCGGCT GCCTGGGCTG 960
 55 AGGGCTCGCT CCAGGGCTTT GCAGACTGGA CCCTTACCGG TGGCTCTTCC TGGCTGGGAC 1020
 CTTCCGCGAG AGTCCCACTA GCCAGCGGCC TCAGCCAGGG ACGAAGGCTT CAAAGCTGAG 1080
 AGGCCCTTAC CGGTGGGTGA TGGATATCAT CCGCGAACAG GTGAAGGGAC AACTGACTAG 1140
 CAGCCCGCAG CCGCTCACCC TGGGATCCCG AGCCTAAAG ACACAGAGA CCTCAGCTAT 1200
 GGAGCCCTTC GGACCACTT CTCACAGACT CTGGCACTGG CCAGGCTCG AACCTGGGAC 1260
 60 CCCTCCTCTG ATGAACACTA CAGTGGCTGA GGCACTAGCC CCCGCCAGG CCCTGTAGGG 1320
 ACAGCATTTG AAGGACACTA ATTGCAGTTG CTTGGTTGAA AGTGCTGTG CTGGAACCTG 1380
 CCTGTACTCA CTATGGGAG CTGGCCCC

Seq ID NO: 275 Protein sequence:

Protein Accession #: NP_003967.1

65 1 11 21 31 41 51
 MELGLGLST LSHCPWPRRQ PALWPTLAL ALLSSVAEAS LGSAPRSPAP RBGPPFVLAS 60
 70 PAGHLPGGRT ARWCSGRARR PPPQSPRPAP PPPAPPSALP RGGRAARAGG PGRSARAAGA 120
 RGCRLRSGLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180
 RPVSQPCRRP TRYAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 276 DNA sequence

Nucleic Acid Accession #: NM_057091.1

Coding sequence: 783-1445

75 1 11 21 31 41 51
 80 ACTGGCCGCT GAGAGAAGAA TCGGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60
 GGACCCCCAA ATCTGCAGT ACCAGCAGTC AGCCGCCCCA CGCAGGGACC GGCTTACCCC 120
 TCGTCCCGCG CCTCACTCA CTTTCTCCCG CCTCGGCCCG GGCTCCCGAG CTCTCTACTT 180
 CGCGTGTCTA CAACTCAAC TCCCGGTTTC CGTGCCTCTC CACCGCTCGA GTTCTCTACT 240
 CTCCATATCC GAGGGGCCCC TCCAGCATTC TACCCCTCTC CCAACCTCGG GGGACCTAGC 300
 85 CRAAGTAGGG GGGACTGGAT CCGACGGGTG GAGCAGCCAG GTGAGCCCCG AAAGGTGGGG 360
 CGGGCAGGG GCGCTCCAG CCCCACCCCG GGATCTGGTG ACGCTGGGGC TGGAATTGA 420
 CACCGGACGG CTGCGCGGCG GGGCAGGAGG CTGCTGAGGG ATGGAGTTGG GCCCGGCCCC 480
 CAGACAAGGC CCGGGGGCTC CGCCAGCAGC AGGTCCCTCG GCGCCAGACC CTGCTGCCA 540

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

CCGGGGCTG GAGCCCCACA CCGAGGGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600
 TAAAAGAGGC ACTGCCAGGT GTACAGTCCT GGGCATGGGC TGTITGAGCT TCGGGGGAGA 660
 GCCCAGCACT GGTCCCCGGA AAGGTGCCTA GAAGAACAAG GTGCAGGACC CCGTGCTGCC 720
 TCAACAGGAG GTGGGGGAA CAGCTCAACA ATGGCTGATG GGCGCTCCTG GTGTGTATAG 780
 AGATGGAAT TGGACTTGGA GGCCTCTCCA CGCTGTCCCA CTGCCCTCGG CTTAGGCGGC 840
 AGCCTGCCCT GTGGCCACCC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCCT 900
 CCTTGGGCTC CGCGCCCGCC AGCCCTGCCC CCGCGGAAGG CCCCCCGCT GTCTTGGCT 960
 CCGCGCGCG CCACCTGCCG GGGGAGCGCA CCGCGCGCTG GTGCAGTGA AGAGCCCGGC 1020
 GGCGCGCGCC GCAGCTTCTT CGGCGCGCGC CCGCGCGGCC TGCACCCCCA TCTGCTCTTC 1080
 CCGCGGGGG CCGCGGGGG CCGGCTGGGG GCCCGGGCAG CCGGCTCGG GCAGGGGGGG 1140
 CGCGGGGCTG CCGGCTGGCG TCGCAGCTGG TCGCGGTGGG CGCGCTCGGC CTGGGCCACC 1200
 GCTCGGAGCA GCTGGTGGCT TTCCGCTTCT GCAGCGGCTC CTGCGCGCGC GCGGCTCTC 1260
 CACACGACCT CAGCTGGCG AGCCTACTGG CGCGCGGGGC CCTGGAGCG CCGCGGGGCT 1320
 CCGGGCGGCT CAGCGAGGCC TGCTGCCGAC CCACGCGCTA CGAAGCGGTC TCCTTCATGG 1380
 ACGTCAACAG CAGCTGGAGA ACCGTGGACT GCCTCTCGGC CACCGCTGCG GGTCTGCTGG 1440
 GCTGAGGGCT CGCTCCAGGG CTTTGCAGAC TGGACCTTTA CCGGTGGCTC TTCTGCGCTG 1500
 GGACCTCTCC CAGAGTCCG ACTAGCCAGC GGCGCTCAGC AGGGACGAAG GCCTCAAGC 1560
 TGAGAGGCC CTACCGGTGG GTGATGGATA TCATCCCGA ACAGGTGAAG GGAACAACGA 1620
 CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG 1680
 CTATGGAGCC CTTGGAGCCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG 1740
 GGACCCCTCC TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT 1800
 AGGGACAGCA TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGCC TGTGCTGGAA 1860
 CTGGCTGTA CTCACTCATG GGAGCTGGCC CC

Seq ID NO: 277 Protein sequence:

Protein Accession #: NP_003967.1

1 11 21 31 41 51
 MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60
 PAGHLPGGRT ARWCSGRARR PPPQPSRALP RGRARAAGG PGRARAAGA 120
 RGCRLRSQIV FVRLALGLHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180
 RFPVQPCRRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 278 DNA sequence

Nucleic Acid Accession #: NM_057160.1

Coding sequence: 1-714

40
 45
 50
 55
 60

1 11 21 31 41 51
 ATGCCCGGCC TGATCTCAGC CCGAGGACAG CCCCCTCCTG AGGTCTCTCC TCCCCAAGCC 60
 CACCTGGGTG CCCTCTTTCT CCCTGAGGCT CCACTTGGTC TCTCCGCGCA GCGTGGCCTG 120
 TGGCCCAACC TGGCCGCTCT GCGCTCTGCT AGCAGCGTCG CAGAGGCCTC CCGTGGCTCC 180
 GCGCCCGCA GCCCTGCGCC CCGCGAAGGC CCCCCGCTG TCCTGGCTCC CCGCGCGGC 240
 CACCTGCCGG GGGGAGCGAC GGGCGGCTGG TGCACTGGAA GAGCCCGCGG GCGCGCGCG 300
 CAGCCTTCTC GGCCCGCGCC CCGCGCGGCT GCACCCCAT CTGCTCTTCC CCGCGGGGG 360
 CGCGCGCGCC GGGCTGGGG CCGGGGAGC CGCGCTCGGG CAGCGGGGGC GCGGGGCTGC 420
 CGCCTGCGCT CGCAGCTGGT GCGGCTCGGC GCGCTCGGCC TGGGCCACCG CTCGACGAG 480
 CTGGTGGGTT TCGCTTCTG CAGCGGCTCC TGCCCGCGCG CCGCTCTTCC ACACGACCTC 540
 AGCCTGGCCA GCCTACTGGG CCGCGGGGCC CTGCGACCGC CCGCGGGCTC CCGCGCGGTC 600
 AGCCAGCCCT GCTGCCGACC CACGCGCTAC GAAGCGGTCT CCTTCATGGA CGTCAACAGC 660
 ACCTGAGAAA CCGTGAGCCG CCTCTCCGCC ACCGCTCGCG GCTGCTGGG CTGAGGGCTC 720
 GCTCCAGGGC TTTGCAGACT GACCCCTTAC CGTGGCTCT TCCTGCTGG GACCCCTCCG 780
 CAGAGTCCCA CTAGCCAGCG GCCTCAGCCA GGGACGAAG CCTCAAAGCT GAGAGGCCCC 840
 TACCGTGGT TGGTGGATAT CATCCCCGAA CAGGTGAAG GACAACTGAC TAGCAGCCCC 900
 AGAGCCCTCA CCTGCGGAT CCGAGCCTAA AAGACACAG AGACCTCAGC TATGGAGCCC 960
 TCGGACCCA CTTCTACAG ACTCTGGCAC TGGCCAGGCC TGAACCTGG GACCCCTCCT 1020
 CTGATGAACA CTACAGTGGC TGAGGCATCA GCGCCCGGCC AGGCCCTGTA GGGACAGCAT 1080
 TTGAAGGACA CATATTGCG TTGCTTGGTT GAAAGTGCT GTGCTGGAAC TGGCTGTATC 1140
 TCACTCATGG GAGCTGGCCC C

Seq ID NO: 279 Protein sequence:

Protein Accession #: NP_476501.1

65
 70

1 11 21 31 41 51
 MPGLISARGQ FLLEVLPPQA HLGALFLPEA FLGLSAQPAL WPTLAALALL SVAEASLGS 60
 APRSPAPREG PPPVLASPAH HLPGGRTARW CSGRRARPPP QPSRPAPPPP APPSALPRGG 120
 RAARAGGPGS RARAAGARG RLRSQLVPVR ALGLGHRSD E LVRFRFCGSGS CRRARSPHDL 180
 SLASLLGAGA LRPPFGSRPV SQPCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG

Seq ID NO: 280 DNA sequence

Nucleic Acid Accession #: NM_057090.1

Coding sequence: 29-715

75
 80
 85

1 11 21 31 41 51
 CTGATGGGCG CTCCTGGTGT TGATAGAGAT GGAACCTGGA CTTGAGGGCC TCTCCAGCT 60
 GTCCCACTGC CCCTGGGCTA GCGGCGAGGC TCACCTTGGT CTCTCGGCGC AGCCTGCCCT 120
 GTGGCCCACT CTGGCGGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCTC CCGTGGGCTC 180
 CGCGCCCGCG AGCCCTGCCA CCGCGAAGG CCCCCGCTG GTCTGGGCT CCGCGCGGC 240
 CCACCTGCCG GGGGAGCGCA CCGCCCGCTG GTGCAGTGA AGAGCCCGGC GCGCGCGGC 300
 GCAGCCTTCT CGGCGCGCG CCGCGCGGCC TGACACCCCA TCTGCTCTTC CCGCGGGGG 360
 CCGCGCGGCG CCGGCTGGGG GCGCGGGCAG CCGCGCTCGG GCAGCGGGG GCGGGGCTG 420
 CCGCTGCGC TCGCAGCTGG TGCGGCTGG CCGCTCGGC CTGGGCCACC GCTCCGAGCA 480
 GCTGGTGGCT TTCCGCTTCT GCAGCGGCTC CTGCGCGGC GCGGCTCTC CACAGCACT 540
 CAGCTGGGCC AGCCTACTGG GCGCGGGGC CCGCGAGGC CCGCGGGCT CCGCGCGGCT 600
 CAGCCAGCCC TGCTGCCGAC CACGCGGCTA CGAAGCGGTC TCCTTCATGG ACGTCAACAG 660

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

CACCTGGAGA ACCGTGGACC GCCTCTCCGC CACGCGCTGC GGCTGCCTGG GCTGAGGGCT 720
CGCTCCAGGG CTTTGCAGAC TGGACCCCTTA CCGTGGCTC TTTCTGCCTG GGACCCCTCC 780
GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGAAGGAG GCCTCAAAGC TGAGAGGCCC 840
CTACCGGTGG GTGATGGATA TCATCCCGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900
CAGAGCCCTC ACCTGCGGA TCCAGCCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960
CTTCGGACCC ACTTCTACA GACTCTGGCA CTGGCCAGGC CTGGAACCTG GGACCCCTCC 1020
TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCGCC CAGGCCCTGT AGGACACGA 1080
TTTGAAGGAC ACATATTGCA GTTGCTTGTG TGAAGTGCC TGTGCTGGAA CTGGCTCTGA 1140
CTACTCATG GGAGCTGGCC CC

Seq ID NO: 281 Protein sequence:
Protein Accession #: NP_476431.1

1 11 21 31 41 51
MELGLGLST LSECPWPRRQ APLGLSAQPA LWPTLAALAL LSSVAEASLG SAPRSPAPRE 60
GPPPVLASPA GHLPGRRTAR WSGRRARRPP PQPSRPAPPP PAPPSALPRG GRAARAGGPG 120
SRARAAGARG CLRLSQLVLE RALGLGHRSD ELVRFRFCGS SCRARSPPHD LSLASLLGAG 180
ALRPPPGSRP VSQPCCRPTR YEAVSFMDVN STWRTVDRLS ATACGCLG

Seq ID NO: 282 DNA sequence
Nucleic Acid Accession #: Eos sequence

1 11 21 31 41 51
CTACTGCACC TGCCCTCTGT TTCCTTTGGA AATCTCTTAC CTTTCATTAG GGTTCCTTTC 60
ATAGCAATTT CCTTTGGTTT TTAAGACTTC TACATTGCTT TTTCTTTTAT TATCTGTGCT 120
CCGTGAACCT TATGAATGCT GCTTAAAAAT AATGTCAAAA TATGTTTTAG CTGCCTACTC 180
AGGTAAAGTT TTTCTTTGCT CTGATCTTGG TTTCCATATA CTATTTTGGG TTTTGTGTA 240
GATCTAATCA ATGATCTAGT CAGAAGCTAC TTCACTGGCT AACAGTGATC ATGTTTCATG 300
GCTAAAAATG AACTTGAAGC ACGGAAGTAG TGGTTGGTCC AGTTTGAAG CTCTTATTAG 360
TATTTCTCAT CCTGGCTGTA ATAATAGCCA TTATTTGTTA TGCCCTTTGT ATGTAGCAGA 420
CACTCTTAAG GATTTTATGT GTATTATTCA AATTGCTATT ACTGTTCTTT TTATAGTTGA 480
GAATCTCAGG ATACCTACAT TTATCACTTT TTCAATATAT ATGTAATTTCT TATT

Seq ID NO: 283 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 564-1481

1 11 21 31 41 51
GAGACTTTTA ATCATCTATC CCTTGTGCTT TAOCGAGACC CTACAATACA CTAGAGGCTT 60
CAAAGAGGTC AAAAATTAC ATGTGTAGAC AAATTAGGTC CCTTAAGATG CCAGGCAAAAC 120
GAAGTGCTAC CAAACACGC AATGACTGTC CTAAGAGTGC GTTCTGGGAT ACACCTGTAA 180
ACTTGGATCA AGTTCCTCTC CCTCTCCTCA AAATATATCG ACTTGTGCTG AAAGAAATCA 240
CGACCGATGC TCACAAATCT GACCTGTGTA TTATATAGGG GGTGGTTTTG GTTCTGCGT 300
CTTTCCTGGA TTCACTGGCA GGTAAACATAT TTCAATGTACA AAATGAACTG CAACACCAAG 360
GCAAAACAGG GACAGGCCCT CAAAGTTGTC GGTAGGAGGC CAGGACCCCG CCAGTGGCGT 420
GGGAGAGCAC CGTACTAAAC AAGCTTGCAA ACAGCAGGCA CCTTCTGCCC ACTGAGGAGG 480
AAGGGCTGCG TAAGGGAGGC CGGGGGCGAG GAAGCCAGC TCTGCAGGCC CTGACAAAGT 540
CCTCCCGGCC TCCACGCGTC GCCATGGCAA CGCGGGGTCT GTGCTGGCCG GGATTTGGCG 600
GCCTGGCGCG CGCAGGGGCC GCTGGGAAAG CGCGTCCCG CGCGGGCTCC GCCAGTTTGA 660
ACTTGGCGGG CCAATGTGGG CGCGGGGGGC GCTGGGGGCC TACTTTTCCC TCTTCTCAAG 720
COGTTTCTC TGCTGACTGC AGACCCAGGT CTCGGCCCTC CTCGGACTCC TGCTCAGTCC 780
CTATGACGGG CGCAGCTGGG CAGGGGCTGG AGGTGGTGGG CTCGCCGTGC CCGCCGCTGC 840
CGCTGAGCTG CAGCAATTC ACCAGTTCGC TGTGTCTCC CCTTGGCCAC CAGAGCTTCC 900
AGTTTGAAGA GGAAGACGCT GACGGGAGG ATGAGGAAGA CGTGATGAT GAGGAAGACG 960
TGGATGAAGA TGCCCATGAT TCAGAGGCCA AAGTGGCGAG CCTGAGAGGA ATGGAGTTAC 1020
AGGGGTGGCC CAGCACTCAG GTTGAATCAG AAAATAACCA AGAAGAACAG AAACAGGTGC 1080
GCTTACCGA AAGCGCCTG ACACCATGGG AGGTGTGGTT TATTGGCAAA GAAAAAGAAG 1140
AAGCTGACCG GCTGCAACTG AAGGCTCTAG AGGAATTAAA TCAACACTA GAAAAAGAA 1200
AAGAAATGGA AGAAGCTGAA AAAAGAAAGA TAATTGCTGA AGAAAAAGCA AAGGAATGGG 1260
TTCAGAAAAA GAATGAGCAA AAAAGAAAAG AAGAGAAACA AAAAATTAAT AAAGAAATGG 1320
AGGAAAAAGC AGCAAAAGAA CTGGAGAAAG AATACTTGCA AGAAAAAGCA AAAGAAAAAT 1380
ATCAAGAATG GTTAAAGAAA AAAAATGCTG AAGAATGTGA GAGGAAGAAG AAAGAAAAAG 1440
AAAACACAGC CAAGCTGAAA TACAGGAGAA AAAGGAAATA GCAGAAAAAA AGTTTCAAGA 1500
ATGGTTGGAA AATGCGAAAC ATAAACCTCG TCCAGCTGCA AAGAGCTATG GTTATGCCAA 1560
TGGAAACCTT ACAGGTTTTT ACAGTGGAAA TTCTATCCA GAACAGCCT TTTAATATCC 1620
AATTCGCTGG AAACCAATTC ATATGCCACC TCCCAAGAA GCTAAGATC TATCAGGAAG 1680
GAAGAGTAAA AGACCTGTGA TAAGTCAGCC ACACAAGTCA TCATCTCTGG TAATTCATAA 1740
AGCCAGGAGC AATCTTTGCC TTGGAATCT GTGCAGAAAT CAAAGATAGC GTATGTGGAA 1800
AATAACATGA TTTTATCTGG AGCTATTAA TTTAAAAATC AGAAATGTT TTTTACTGCT 1860
CAGTCAATAA CTCACACTT AATGTGATTA TTGACAAATA GCAATTTTGT CATTTGTATA 1920
TGGAGTCTCT AGAGTTGAGG AAGATATTTT CTGGATTTTG GTTTTATATA ACTTTTAAAG 1980
GTTGATCTTG GCATGTTGTT TTGCAGAAAT AGTGGCTGAA TATGTAAGAA TTGTGTTTGT 2040
ATTAGCTTGG TATTAAAGT ACACCTGTAAT ACCAATAAAA CTAACAAATT TTCTTG

Seq ID NO: 284 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MATRGLCWPG LAGLARAGPA GKARPRGSA SLNLAQMWNA AGRNGPTFPS SYAGFSADCR 60
PRSRPSSDSC SVPMTGARGQ GLEVVRSPPS FLPLSCSNST RSLSPGLHQ SFQFDEDDGD 120
GEDEEDVDDE EDVDEDAHDS EAKVASLRGM ELQGCASQV ESENNQEEQK QVRLPESRLT 180
PWEVWFIGKE KEERDLRLQK ALEELNQOLE KRKEMEEREK RKIIAEKHK EKVQKNEQK 240
RKEREQKINK EMEKAAKEL EKEYLQEKAK EKYQEWLKKK NAECEERKK EKVQNSKLKY 300

RRKRX

Seq ID NO: 285 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1-1746

5

1	11	21	31	41	51	
ATGCCACTGA	AGCATTATCT	CCCTTTGCTG	GTGGGCTGCC	AAGCCTGGGG	TGCAGGGTTG	60
GCCTACCATG	GCTGCCCTAG	CGAGTGTACC	TGCTCCAGGG	CCTCCCAGGT	GGAGTGCACC	120
GGGGCAACGA	TTGTGGCGGT	GCCCCACCC	CTGCCCTGGA	ACGCCATGAG	CCTGCAGATC	180
CTCAACACGC	ACATCACTGA	ACTCAATGAG	TCCCCGTTCC	TCAATATCTC	AGCCCTCATC	240
GCCTCAGGGA	TTGAGAAGAA	TGAGCTGTGC	GCATCAACGC	CTGGGGCCTT	CGAAACCTCG	300
GGCTCGCTGC	GCTATCTCAG	CCTCGCCAAC	AACAGCTGTC	AGGTTCTGCC	CATCGGCTTC	360
TTCCAGGGCC	TGGACAGCCT	TGAGTCTCTC	CTTCTGTCCA	GTAACAGCT	GTTGCAGATC	420
CAGCCGGCCC	ACTTCTCCCA	GTGCAGCAAC	CTCAAGGAGC	TGCAGTTGCA	CGGCAACAC	480
CTGGAATACA	TCCCTGACGG	AGCCTTCGAC	CACCTGGTAG	GACTCAGGAA	GCTCAATCTG	540
GGCAAGAATA	GCCTCACCCA	CATCTCACCC	AGGGTCTTCC	AGCACCTGGG	CAATCTCCAG	600
GTCTCTCGGC	TGTATGAGAA	CAGGCTCAGC	GATATCCCCA	TGGCACTTT	TGATGGGCTT	660
GTTAACCTGC	ATGAACCTGC	TCTACAGCAG	AACAGATTTG	GACTGCTCTC	CCCTGGTCTC	720
TTCCACAACA	ACCACAACTC	CCAGAGACTC	TACCTGTCCA	ACAACCATAT	CTCCAGCTCG	780
CCACCCAGCA	TCTTCATGCA	GCTGCCCCAG	CTCAACCGTC	TTACTCTCTT	TGGGAATTCC	840
CTGAAGGAGC	TCTCTCTGGG	GATCTTCGGG	CCCATGCCCA	ACCTGCGGGA	GCTTTGGCTC	900
TATGACAACC	ACATCTCTTC	TCTACCCGAC	AATGTCTTCA	GCAACCTCCG	CCAGTTGCAG	960
GTCTCTGATC	TTAGCCGCAA	TCAGATCAGC	TTCATCTCCC	CGGGTGGCTT	CAACGGGCTA	1020
ACGGAGCTTC	GGGAGCTGTC	CCTCCACACC	AACGCACTGC	AGGACCTGGA	CGGGAATGTC	1080
TTCCGATGAT	TGGCCAACTC	GCAGAACATC	TCCCTGCAGA	ACAATCGCCT	CAGACAGCTC	1140
CCAGGGAATA	TCTTCGCCAA	CCTCAATGGC	CTCATGGCCA	TCCAGCTGCA	GAACAACCA	1200
CTGGAGAATC	TGCCCCCTCG	CATCTTCGAT	CACCTGGGGA	AACGTGTGTA	GCTGGCGCTG	1260
TATGACAATC	CTTGGAGGTC	TGACTCAGAC	ATCCTTCCGC	TCCGCAACTG	GCTCCTGCTC	1320
AACCAAGCTA	GGTATGGGAG	GGACACTGTA	CCTGTGTGTT	TCAGCCAGC	CAATGTCCGA	1380
GGCCAGTCCC	TCATTATCAT	CAATGTCAAC	GTGTCTGTTT	CAAGGCTCCA	TGTCCTGAG	1440
GTGCCTAGTT	ACCCAGAAAC	ACCATGGTAC	CCAGACACAC	CCAGTTACCC	TGACACCACA	1500
TCCGTCTCTT	CTACCACTGA	GCTAACCAAG	CCTGTGGGAG	ACTACACTGA	TCTGACTACC	1560
ATTCAAGTAT	CTGATGACCG	CAGCGTTTGG	GGCATGACCC	AGGCCAGAG	CGGGCTGGCC	1620
ATTGCCGCCA	TTGTAATGGG	CAATGTGGCC	CTGGCCTGCT	CCCTGGCTGC	CTGGCTGGCC	1680
TGTGTCTGCT	GCAGAAGAG	GAGCCAAGCT	GTCTGTATGC	AGATGAAGGC	ACCCAATGAG	1740
TGTTAAGAG	GCAGGCTGGA	GCAGGCTGCG	GGAAATGATG	GACTGGAGGA	CCTGGGAATT	1800
TCATCTTTCT	GCCTCCACCC	CTGGGTCCAT	GGAGCTTTCC	CGTGATTGCT	CTTTCTGGCC	1860
CTAGATAAAG	GTGTGCTCAG	CTCTTCTGTA	CTTGCCTGAT	TCTCCCGTAG	AGAAGCAGGT	1920
CGTGCCGGAC	CTTCTTACAA	TCAGGAAGAT	AGATCCAACT	GGCCATGGCA	AAAGCCCTGG	1980
GGATTTCCGA	TTCATACCCC	TGGGCTTCTC	TCGAGAGGGC	TCTTCTCCCA	AATCTCCCCC	2040
ACCTGTCTCT	CAAGAACAGC	CTTCCCTGCG	CCCAGGCCCC	CTCCGGGCTC	CTGTAGACTC	2100
AGTTAGTCCA	CAGCCTGCTC	ACTTCGTGGG	AATAGTTTCT	CGCTGAGATA	GGCCCTCTCG	2160
CCCTAAGTAT	ATGTAAGTTG	ATTTCCCTTC	TTTGTCTTCT	CTTGTGTTGT	CTATGGCTTG	2220
ACCCAGCATG	TCCCTCAAA	TGAAAGTTCT	CCCTTGATT	TTCTGCTCCT	GAAGGCAGGG	2280
TGAGTTCTCT	CCTCAAGAA	GACTTCAAAC	CATTTAATG	GTTTCTTAAG	AGCCGTCAAT	2340
CAGCCTGGTT	TTGGGGATGC	TATGAAAGAG	AGAAGGAAAA	TCATGCCGCT	CAGTTCTCTG	2400
AGACAGAAGA	CGCCTCATCA	GTGTCTCACT	TGTGATTTTT	ATCTGGAAAA	GGAGAAACAA	2460
CCCCAGCACA	CCAAAGCTCAG	CCTTTTAGAG	AAGGATATTT	CCAACTGCGA	AACCTTGCTT	2520
TGAAAGTTT	AGCCCTTTAA	GGAAATGAAT	CATGTAGAAT	TTTGGACTTC	TAAAAACATT	2580
AAAATCAGCT	TATTAATACG	GGATAGAGAA	AGAAATCTGG	TGCTTGGGGG	TCCCTGTGTT	2640
CACCCCTAGA	GTTTGTTTTA	AAATTTTTAA	TTGAAGCATG	TGAAGTGATC	STGCAGAAAA	2700
GTGGGAACAT	GATAGTGTAT	GGCTTGGTGG	ATTTTCACAA	ACTGAACATA	CCTGTGTAAT	2760
CAGCATCTAG	ACCCAGCTCAG	AGAGCATCAC	AAATATCCCC	CATCCTGGGC	TTTTCCAGAA	2820
GGAGATGGGG	GCTTCTGAAG	ATGGACTTAC	CTGGGACCTG	CCCCCATGAA	GCCAGGACGG	2880
TCCCCCACA	GTCAGCTGT	GCAAAGGCCC	CGTGGCCAGG	GGTGGAGGAG	AATATGTGGG	2940
TGTGGACAGG	ATGGGAGATC	GTGGCCTGAA	CAGGAGATTT	TATTATATCT	GGAGACCTCG	3000
AGAGACCCCT	AGACCTGGGG	CACCATGGCT	GGCCAGGTCA	GAAGCATCCT	GACTGCAGAG	3060
GTCCGTGCGC	CCACACCTTC	TTCCCTGCCA	GCAAGTTGTC	TGCGGCTCAT	CGGAGGCCCC	3120
TCCGCTGGA	GCCTTCTATG	GACGTGATAT	GCCTGTATCT	GTTTTTAATT	TCAATCTTTC	3180
ACTTAGGGGA	AGTGAATACG	CTCAGAGATG	AGATCCTTTA	ATTGAAAAAG	AAGTGTAAAG	3240
GAATCTAGTG	TCCTTCTAAT	GTGGTAAAT	TCTCCATCAA	CATCACAGTC	AGCTGGCAGC	3300
TGAATTTAG	AATCTCACTT	ACAGCAGGCG	ACACGGGGGT	ACACCGATGG	GTCACACTGG	3360
GTCTGGGGGC	TCCCTGGAGC	TCCTCTGGCG	TGTGGTCTGG	TTAGGAGTTG	AGTTGTTTGC	3420
TCCAGGGTTA	TCTCTCTCTT	CGAGTCACAG	TCACACGAAT	ACCTGCCTTC	TCTGGCTTTC	3480
CTGCTATACA	CATATTACAA	TGGCGCTCAA	GAAGTTAGGC	TCATGGCAAC	GTGTGTCTTT	3540
CTCTGGACAA	CTGCCCCAGT	TTACAGTGAA	ATGGAGAAAT	TCAGGTCTCC	ACGTCTGCC	3600
AGGAAAGAAC	TTCAAGTGAC	TCCACGGGGA	TCTGGAAATC	CACGACCAAT	CCCGATCGGC	3660
TCCTATTAGC	TCCCCGCTCC	ACAAGACACC	TGTGCTTTGG	AAATCCACCA	CCAATCCCGA	3720
TCCGCTCTTA	TTAGCTCCCC	GCTCCACAAG	ACACCTGTGA	TCTGGAAATC	TACCAACCAAT	3780
CCCGATCGGC	TCCTATTAGC	TCCCGCTCC	ACAAGACACC	TGTGACATCC	TCCAGGGCCA	3840
CAGGAGCAGC	TGCTGACCA	TTTCCCTTC	CAGTTCTGTC	ACAAAAAGTG	TCCAGAGGGC	3900
TGTTTGCAAA	CACATAGTGA	CTTTGTAGCT	TTTCAACCTC	TGTCCAGGGG	AATCTAGGAG	3960
AGATGAGGCG	CGTCAGAGTC	AAGAGATGTC	ATCCCCCAG	GGTCTCCAAG	GCAATTCAC	4020
ACTATTGGTG	GCACCTGGAG	GACATGCACC	AAGGCTTGCC	AGAGCCACAA	GGAAGTGAGC	4080
CCAGAGCATG	GCACATGAGC	ATCACCCTGC	GATGGTGGCC	TGCTGTGCTT	GGTGCCAACA	4140
GGGGCATCCC	GGCCCGTACC	CCTCCAGACA	GGAAGCATGG	GTTCGCCCA	AGACCTGTGG	4200
GGTGCTCCTG	TGAGTGGCCT	CCAGATGTCT	TTGTGCATAG	GCACAAGTGG	GCCAGGGCTG	4260
GAGGGAGGTG	GGAAACCTCA	TCATCCGGTG	GGCCCTGCCA	ATCTTAACCC	AGAACCTCTA	4320
GGTATTCTCG	GCAGTAGCCA	TGACATTGGA	GCACCTTCTC	CTCCAGCCAG	AGGCTGACCT	4380
GAGGGCCACT	GTCTCAGAT	GACACCAACC	AGGAGCACCC	TAGGTGAGGG	GTGAGGGCCC	4440
CCTTATGTGA	ACCTCTTGGC	TCTTCTTTTC	TCCCATCAGA	GTGGTTGGAT	GGAGCCATTG	4500
GCCTCCTTTT	CTTCAGGGGG	CCCTTCAACC	TCTCTGCACC	ATGTTGTCTG	GCTGAGGAGC	4560
TACTAGAAAA	GCTGAGTGGG	GTCTCCTTTC	CAACAGGATG	ATGCATTGTC	TCAATTCTCA	4620
GGGCTGGAAAT	GAGCCGGCTG	GTCCCCCAGA	AAGCTGGAGT	GGGTACAGAA	GTTCAAGTTT	4680
CCTCTCTGTT	TACAGCTCCT	TGACAGTCCC	ACGCCCATCT	GGAGTGGGAG	CTGGGAGTTA	4740

GTGTTGGAGA AGAAACAACA AAAGCCAATT AGAACCCTA TTTTAAAAA GTGCTTACTG 4800
 TGCACAGATA CTCCTCAAGC ACTGGACGTG GATTCTCTCT CTAGCCCTCA GCACCCCTGC 4860
 GGTAGGAGTG CCGCCTCTAC CCACCTTGTA TGGGGTACAG AGGCACCTGC TCTTCTGCAT 4920
 5 GGTGTTCAAT AGGCTGGGAG TTTTATTTAT CTCCTCAAAC TTTGTACAAG AGCTCATGGC 4980
 TTGCTCTGGG CTTTCTGTCAT TAAACCAAG GAAATGGAAG CCATTCCCTT GTTGTCTCTCC 5040
 TTAGTCTTGG TCATCAGAAC CTCACCTGGT ACCATATAGA TCAAAAGCTT TGTAAACCACA 5100
 GGAATAAATA AACTCTTCCA TCCCTTAAAG AATAGAAATG TTTGTCCCTC TCATGGGAAT 5160
 TGGGCTGTAT GTATATTGTT CTCTCTCTCT AGAATTTAGA GATACAAGAG TTCTACTTAG 5220
 10 AACTTTTCAT GACACAAAT TCCACAACT TTCAGATGCT GATGTAGAGC TATTGGGAAA 5280
 GAACTTCCAA ACTCAGGAAG TTTGCAGAGA GCAGACAGCT AGAGATAACT CGGGACCCAG 5340
 AGTTGGTCCA CAGATGTTAG ATGTATCCTA GCTTTTAGCC ATAAACCACT CAAAGATTCA 5400
 GCCCCAGAT CCCACAGTCA GAACTGAATC TGGTGTGTTG GGAAGCCAGC AGTGGCCTTG 5460
 GGAAGGAAGC CATGGCTGTG GTTCAGAGAG GGTGGGCTGG CAAGCCACTT CCGGGGAAAA 5520
 15 CTCCTTCCGC CCCAGGTTTC TTCTTCTCTT AAGGAGAGAT TGTCTCTACC AACCCGCTGC 5580
 CTTATGCTCG CTCCTCAAAGC TAGATCATGT TTGCTTGTCT TAGAGAATTA CTGCAAAATCA 5640
 GCCCCAGTGC TTGGCGATGC ATTTACAGAT TTCTAGGCC TCAGGGTTT GTAGAGTGTG 5700
 AGCCCTGGTG GGCAGGGTTG GGGGGTCTGT CTTCTGCTGG ATGCTGCTTG TAATCCATT 5760
 GGTGTACAGA ATCAACAATA AATAATATAC ATGTAT

Seq ID NO: 286 Protein sequence:
 Protein Accession #: NP_570843.1

1 11 21 31 41 51
 25 MPLKHYLLLL VGCQAWGAGL AYHGCPSPECT CSRASQVECT GARIVAVPTP LFWNAMSLLQI 60
 LNTHTLNLNE SPFLNISALI ALRIEKVELS RITPGAFRNL GSLRYLSLAN NKLQVLPGL 120
 FQGLDSLESLL LLSNQLLQI QPAHFSQCSN LKELQLHGNH LEYIPDGAFF HLVLGLTKNL 180
 GKNSLTHISP RVPHQLNLQ VLRLYENRLT DIPMCTFDGL VNLQELALQQ NQIGLLSPGL 240
 30 FHNHNLRLQL YLSNNHISQL PPSIFMQLPQ LNRLLTLPNS LKELSLGIFG PMPNRLRELWL 300
 YDNHISLLPD NVFSNLRQLQ VLILSRNQIS FISPAGFNFL TELRELSLHT NALQDLGNV 360
 FRMLANLQNI SLQNNRLRLQ PGNIFANVNG LMAIQLQNNQ LENLPLGIFD HLGKLCLERL 420
 YDNPWRCDSD ILPLRNWLLL NQPRLGTDV FVCFSPANVR GQSLIIINVN VAVPSVHVPE 480
 VPSYPETPHY PDTPSYPDIT SVSSTELTS PVEDYTDLIT IQVTDDRSVW GMTQAQSGLA 540
 35 IAAIVIGIVA LACSLAACVG CCCCKKRSQA VLMQMKAPNE C

Seq ID NO: 287 DNA sequence
 Nucleic Acid Accession #: NM_002362
 Coding sequence: 1..954

1 11 21 31 41 51
 40 ATGTCITCTG AGCAGAAGAG TCAGCACTGC AAGCCTGAGG AAGGCGTGA GGGCCCAAGAA 60
 GAGGCCCTGG GCCTGGTGGG TGCACAGGCT CCTACTACTG AGGAGCAGGA GGCTGCTGTC 120
 45 TCCTCTCTCT CTCTCTCTGT CCCTGGCACC CTGGAGGAAG TGCCCTGCTGC TGAGTCAGCA 180
 GGTCCTCCCC AGAGTCTCTA GGGAGCCTCT GCCTTAACCA CTACCATCAG CTTCACTTGC 240
 TGGAGGCAAC CCATAGAGGG TTCCAGCAGC CAAGAAGAGG AGGGGCCAAG CACCTCGCCT 300
 GAGCAGAGT CCTTGTTCAG AGAAGCACTC AGTAACAAGG TGGATGAGTT GGCTCATTTT 360
 CTGCTCCGCA AGTATCGAGC CAAGGAGCTG GTCACAAAGG CAGAAATGCT GGAGAGAGTC 420
 50 ATCAAAATTT ACAAGCGCTG CTTTCTCTGT ATCTTCGGCA AAGCCTCCGA GTCCCTGAAG 480
 ATGATCTTGT GCATTGACGT GAAGGAAGTG GACCCCGCCA GCAACACCTA CACCCTTGTC 540
 ACCTGCTGGG GCCTTTCCTA TGATGGCCTG CTGGGTAAATA ATCAGATCTT TCCCAAGACA 600
 GGCCTTCTGA TAATCGTCTT GGGCACAATT GCAATGGAGG GCGACAGCGC CTTCTAGGAG 660
 GAAATCTGGG AGGAGCTGGG TGTGATGGG GTGTATGATG GGAGGGAGCA CACTGTCTAT 720
 55 GGGGAGCCCA GGAACCTGCT CACCCAAGAT TGGGTGACAG AAAACTACCT GGAGTACCGG 780
 CAGGTACCCG GCAGTAATCC TGCGCGCTAT GAGTTCCTGT GGGGTCCAAG GGCTCTGGCT 840
 GAAACAGCT ATGTGAAAGT CCTGGAGCAT GTGGTCAGGG TCAATGCAAG AGTTCGCATT 900
 GCCTACCCAT CCCTGCGTGA AGCAGCTTTG TTAGAGGAGG AAGAGGGAGT CTGA

Seq ID NO: 288 Protein sequence:
 Protein Accession #: NP_002353.1

1 11 21 31 41 51
 65 MSSEKQSQHC KPEEGVEAQE EALGLVGAQA PTTEEQEAHV SSSSPLVPGT LEEVPAESA 60
 GPPOSQOAS ALPTTISFTC WRQPNBSSS QEEBGPSTSP DAESLFREAL SNKVDELAHF 120
 LLRKYRAKEL VTKAEMLERV IKNYKRCFPV IFGKASESLK MIFGIDVKEV DPASNTYTLV 180
 TGLGLSYDGL LGNNQIFPPT GLLIIVLGTI AMEGDSASEE EIWEELGVMG VYDGREHTVY 240
 70 GEPRKLLTQD WVQENYLEYR QVPGSNPARY EFLWGPRLA ETSYVKVLEH VVRVNRVRI 300
 AYPRLREAL LEEEGV

Seq ID NO: 289 DNA sequence
 Nucleic Acid Accession #: NM_002362
 Coding sequence: 46..1344

1 11 21 31 41 51
 80 CGGCGGCCGC GCCTGGTGG GGTCCCACT GCTCTCGGGG GCGCCATGGA CGAGGCCGTG 60
 GCGACCTGA AGCAGGCGCT TCCCTGTGTG GCGAGTGC CAACGGTCCA CGTGGAGGTG 120
 CATCAGCGCG GCACAGCAC TGCAAGAAA GAAGACATAA ACCTGAGTGT TAGAAGCTA 180
 CTCACAGAG ATAAATATTG GTTTGGTGAT TACACATGGA CTGAGTTTGA TGAACCTTTT 240
 TTGACAGAA ATGTGCAGTC TGTGTCTATT ATTGACACAG AATTAAAGST TAAAGACTCA 300
 85 CAGCCCATCG ATTTGAGTGC ATGCACTGTT GCACCTTACA TTTCCAGCT GAATGAAGAT 360
 GGCCCCAGCA GTGAAATCT GGAGGAAGAG ACAGAAAACA TAATTGCAGC AAATCACTGG 420
 GTTCTACCTG CAGCTGAATT CCATGGGCTT TGGGACAGCT TGGTATACGA TGTGGAAGTC 480
 AAATCCCATC TCCTCGATTA TGTGATGACA ACTTTACTGT TTTGACACA GAACGTCAAC 540

AGCAACCTCA TCACCTGGAA CGGGTGGTG CTGCTCCAGG GTCCTCTGG CACTGGAAAA 600
 ACATCCCTGT GTAAAGCGTT AGCCAGAGAA TTGACAATTG GACTTTCAG CAGGTACCGA 660
 TATGGCCAAT TAATTGAAAT AAACAGCCAC AGCCTCTTTT CTAAGTGGTT TTCGGAAGT 720
 GGCAAGCTGG TAACCAAGAT GTTTCAGAAG ATTGAGATT TGATTGATGA TAAAGACGCC 780
 CTGCTGTTCG TGCTGATTGA TGAGGTGGAG AGTCTCACAG CGCCCGGAAA TGCCTGCAGG 840
 GCGGGCACCG AGCCATCAGA TGCCATCCGC GTGGTCAATG CTGCTTGAC CCAAAATTGAT 900
 CAGATTAAAA GGCATTCCAA TGTGTGATT CTGACCATT CTAACATCAC CGAGAAGATC 960
 GACGTGGCCT TCGTGGAGCA GGCTGACATC AAGCAGTACA TTGGGCCACC CTCTGCAGCA 1020
 GCCATCTTCA AAATCTACCT CTCTGTGTTG GAAGAAGTGA TGAAGTGCA GATCATATAC 1080
 CCTCGCCAGC AGCTGCTGAC CCTCGAGAG CTAGAGATGA TTGCTTCAT TGAAGAACAC 1140
 GTGTCAAAAT TGAGCCTTCT TTGGAATGAC ATTTCAAGGA AGAGCGAGGG CCTCAGCGGC 1200
 CGGGTCTCTG AAAAACTCCC CTTTCTGGCT CATGCGCTGT ATGTCAGGC CCCACCGTC 1260
 ACCATAGAGG GGTCTCTCCA GGCCCTGTCT CTGGCAGTGG ACAAGCAGTT TGAAGAGAGA 1320
 AAGAAGCTTG CAGCTTACAT CTGATCCTGG GCTTCCCAT CTGGTGTCTT TCCATGGAG 1380
 AACACACAAC CAGTAAGTGA GGTGCCCCA CACAGCGTC TCCAGGGAA TCCCTTCTGC 1440
 AAACCAACG TTACTTAGAG TGCAAGCTAG AAAGCCACCA AGGCCAGGCT TTGTTAAAG 1500
 AAGTGATTCT TATTTATGTT GTTTTAAAT GCATCTGAG AGACAAACAT CTGTGCTATT 1560
 TCACGTGTTG TAAAGATAA TTCAGATTGT TTGTCTCCT GTGAAGAAC ATCGAAACCT 1620
 GTTTGTTCCT AGCCACCCCT CAGTGGATGG GATGCATAAT GCCAGCAAGT TTTGTTTAA 1680
 AGCAAAAAAG GAAGATTAA GCAGGTGTTA TAGAAGCCAG AAGAGAACT GTGTACCCCT 1740
 AAAGAAGCAT ATATCATAG CATTAAAAAT GCACACATTA CTCCAGGTGG AAGGTGGCAA 1800
 TTGCTTTCTG ATATCAGCTC GTTTGATTGA GTGCAAAAT GTTTTCAAGA CTATTTAATG 1860
 GATGTAAAAA AGCCTATTTC TACATTATAC CAACTGAGAA AAAAAAGTGC GGTAAAGTGT 1920
 TCTTTTATAA TAAATAATCA AGACATGCTC CCATTGTCAG GAAAAGTGCA GACTCTGAGT 1980
 GTTCCAGGGA AACACATGCT GGACATCCCT TGTAAACCGG TATGGCGGCC CCTGCATTGC 2040
 TGGGATGTTT CTGCCCCAGG TTTGTGTTGT GCAATAACGT TATCATATT CTAAATGAGG 2100
 TTCACATTA TATAATATAA AATAAATAGG TCAGTTACTG GTCTCTTCT GCGGAATGTT 2160
 ATGTTTGTCT TTTATCTCAC AGTAAATAA ATATAATTAA AAA

30 Seq ID NO: 290 Protein sequence:
 Protein Accession #: NP_004228

1 11 21 31 41 51
 35 MDEAVGLKQ ALPCAESPT VEVEVHQRGS STAKKEDINL SVRKLNRHN IVPGDYTWTE 60
 FDEPFLTRNV QSVSIIDTEL KVKDSQPIDL SACTVALHIF QLNEDGPSSE NLEETEENII 120
 AANHVLPAE EFHGLWDSL VYDEVKSHLL DYVMTLLFS DKNVNSNLI WNRVLLHGP 180
 PGTGKSLCK ALAQKLTRL SSRYRYGQLI EINSLSLFSK WFSSEGLVLT KMFQIKDLI 240
 40 DDKDALFVL IDEVESLTA RNACRAGTEP SDAIRVNAV LTQIDQIKRH SNVVLITTSN 300
 ITERKIDVAFV CRADGKYIG PPSAAAI FKI YLSLEELMK QIIYFPRQL LTLRELEMIG 360
 FIENNVSKLS LLLNDISRKS EGLSGRVLRK LPFLAHALYV QAPTVTIEGF LQALS LAVDK 420
 QFEERKLAA YI

45 Seq ID NO: 291 DNA sequence
 Nucleic Acid Accession #: NM_002658.1
 Coding sequence: 77-1372

1 11 21 31 41 51
 50 GTCCCGCAG CGCGTCCGG CCTCCTGCC GCAGGCCACC GAGGCGCGCG CGGTCTAGCG 60
 CCCGAGCTC GCCACATGA GAGCCCTGCT GCGCGCGCTG CTCTCTGCG TCTCTGGTCT 120
 GAGCGACTCC AAGAGCAGCA ATGAACCTCA TCAAGTTCCA TCGAACTGTG ACTGTCTAAA 180
 TGGAGGAACA TGTGTGTCCA ACAAGTACTT CTCCAACATT CACTGGTGCA ACTGCCCAA 240
 55 GAAATTCGGA GGGCAGCACT GTGAATAGA TAAGTCAAAA ACCTGCTATG AGGGGAATGG 300
 TCACTTTAC CGAGGAAAGG CCAGCACTGA CACCATGGGC CGGCCCTGCC TGCCCTGGAA 360
 CTCTGCCACT GTCCCTCAGC AAACGTACCA TGCCCAAGCA TCTGATGCTC TTCAGCTGGG 420
 CCTGGGAAA CATAATTACT GCAGGAACCC AGACAACCGG AGGCGACCTT GGTGCTATGT 480
 GCAGGTGGGC CTAAGCCGCT TTGTCCAAGA GTGCATGGTG CATGACTCGG CAGATGGAAA 540
 60 AAAGCCCTCC TCTCTCCAG AAGAATTAAA ATTTCACTGT GGCCAAAAGA CTCTGAGGCC 600
 CGCCTTTAAG ATTAATTGGG GAGAATTAC CACCATCGAG AACCCAGCCT GGTTCGCGC 660
 CATCTACAGG AGGCACCGGG GGGGCTCTGT CACCTACGTG TGTGGAGGCA GCCTCATCAG 720
 CCCTTCTGG GTGATCAGCG CCACACACTG CTTCAATTGAT TACCCAAAAG AGGAGGACTA 780
 CATGCTCTAC CTGGGTCCGT CAAGGCTTAA CTCCAACAG CAAGGGGAGA TGAAGTTTGA 840
 GGTGGAAAA CTATCTCTAC ACAAGGACTA CAGCGCTGAC ACCTTGTCTC ACCACAACGA 900
 65 CATTGCTTGG CTGAAGATCC GTTCCAAGGA GGGCAGGTGT GCGCAGCCAT CCGGACTAT 960
 ACAGACCATC TGCCCTGCCCT CGATGTATAA CGATCCCGAG TTTGGCACA GCTGTGAGAT 1020
 CACTGGCTTT GGAAAAGAGA ATTCTACCGA CTATCTCTAT CCGGAGCAGC TGAAGATGAC 1080
 TGTGTGAAG CTGATTTCCT ACCGGGAGTG TCAGCAGCCC CACTACTACG GCTCTGAAGT 1140
 70 CACCAACAAA ATGCTATGTT CTGCTGACCC CCAATGGAAA ACAGATTCTT GCCAGGGAGA 1200
 CTCAGGGGGA CCCCTGCTCT GTTCCCTCCA AGGCCGCTG ACTTTGACTG GAATTGTGAG 1260
 CTGGGGCCGT GGATGTGCC TGAAGGACAA GCCAGGCGTC TACACGAGAG TCTCACACTT 1320
 CTTACCTTGG ATCCGAGTCT ACACCAAGGA AGAGATGGC CTGGCCCTCT GAGGTCCCC 1380
 AGGGAGGAAA CGGCACCAAC CCGCTTTCTT GCTGGTTGTC ATTTTTCAG TAGAGTCATC 1440
 75 TCCATCAGCT GTAAGAGAG ACTGGGAAGA TAGGCTCTGC ACAGATGGAT TTGCTGTGG 1500
 CACCACCAAG GTGAACGACA ATAGCTTTAC CCTCACGGAT AGGCCTGGGT GCTGGCTGCC 1560
 CAGACCTCTT GGCACAGGAT GAGGGGTGGT CTGACTCAA CATGTTACTG ACCAGCAACT 1620
 TGTCTTTTTC TGGACTGAAG CCTGCAGGAG TTAAAAAGGG CAGGGCATCT CTTGTGCATG 1680
 GGCTCGAAGG GAGAGCCAGC TCCCCGACC GGTGGGCATT TGTGAGGCCC ATGGTTGAGA 1740
 80 AATGAATAAT TTCCCAATTA GGAAGGTAA GCAGCTGAGG TCTCTTGAGG GAGCTTAGCC 1800
 AATGTGGGAG CAGCGGTTTG GGGAGCAGAG ACACTAAGCA CTTGAGGGCA GGGCTCTGAT 1860
 ATTCCATGAA TGTATCAGGA AATATATATG TGTGTGATG TTTGCACACT TGTGTGTGG 1920
 GCTGTGAGTG TAAGTGTGAG TAAGAGCTGG TGTCTGATT TTAAGTCTAA ATATTTCCTT 1980
 85 AAACCTGTGT GAGCTGTGAT CCACACAGAG TGGTCTTTCT GGAGAGGTTA TAGGTCATC 2040
 CTGGGGCCTC TTGGGTCCTC CAGCTGACAG TGCCCTGGGA TGTACTTATT CTGCAGCATG 2100
 ACCTGTGACC AGCATGTCT CAGTTTCACT TTCACATAGA TGTCCTTTC TTGGCCAGTT 2160
 ATCCCTTCCT TTTAGCTTAG TTCAATCCAT CCTCACTGGG TGGGGTGAGG ACCACTCCTT 2220
 AACTGAATA TTTATTTTCT ACTATTTTAA TTTATATTTT TGAATTTTAA AATAAAAGTG 2280

ATCAATAAAA TGTGATTTT CTGA

Seq ID NO: 292 Protein sequence:
Protein Accession #: NP_002649.1

5

1	11	21	31	41	51	
MRALLARLLL	CVLVVSDDSG	SNELHQVPSN	CDCLNGGTCV	SNKYFSNIHW	CNCPKFGGGQ	60
HCEIDKSKTC	YEGNGHFYRG	KASTDTMGRP	CLPWN SATVL	QQTYHAHRSD	ALQLGLGKHN	120
YCRNPDNRRR	PWCYVQVGLK	PLVQECMVHD	CADGKKPSSP	PEELKFCQCG	KTLLRPRFKII	180
GGFTTITENQ	PWFAAIYRRH	RGGSVTYVCG	GSLISPCWVI	SATHCPIIDYF	KKEDYIVYLG	240
RSRLNSNTQG	EMKFEVENLI	LHKDYSADTL	AHRNDIALLK	IRSKEGRCQA	PSRTIQITICL	300
PSMYNDPQFG	TSCEITGFGK	ENSTDYLYPE	QLKMTVVKLI	SHRECCQPHY	YGEVTTKML	360
CAADPQWKTD	SCQDGGGGL	VCSLQGRMTL	TGIVSWGRGC	ALKDKPGVYT	RVSHPLEWIR	420
SHTKEENGLA	L					

15

Seq ID NO: 293 DNA sequence
Nucleic Acid Accession #: NM_001498
Coding sequence: 93..2006

20

1	11	21	31	41	51	
GGCAGGAGGC	TGAGTGTCCG	TCTCGCGCCC	GGAAGCGGGC	GACCGCGGTC	AGCCCGGAGG	60
AGGAGGAGGA	GGAGGAGGAG	GAGGGGGCGG	CCATGGGGCT	GCTGTCCAG	GGCTCGCGCG	120
TGAGCTGGGA	GGAAACCAAG	CGCCATGCCG	ACCACGTGCG	GCGGCACGGG	ATCCTCCAGT	180
TCTGTGACAT	CTACCAACGC	GTCAAGGACC	GGCACAAGGA	CGTTCTCAAG	TGGGGCGATG	240
AGGTGAATA	CAGTTGGTGA	TCTTTTGATC	ATGAAAATAA	AAAAGTCCGG	TGTGCTCTGT	300
CTGGGGAGAA	AGTTCCTGAA	ACTCTGCAAG	AGAAGGGGGA	AAGGACAAAC	CCAAACCATC	360
CTACCCCTTG	GAGACACAGAG	TATGGGAGTT	ACATGATTGA	AGGGACACCA	GGACAGCCCT	420
ACGGAGGAAC	AATGTCCGAG	TTCAATACAG	TTGAGGCCAA	CATGCGAAAA	CGCCGGAAGG	480
AGGCTACTTC	TATATTAGAA	GAAAATCAGG	CTCTTTGCAC	AATAACTTCA	TTTCCAGAT	540
TAGGCTGTCC	TGGGTTTACA	CTGCCCGAGG	TCAAACCCAA	CCCACTGGAA	GGAGGAGCTT	600
CCAAGTCCCT	CTTCTTTCCA	GATGAAGCAA	TAAACAAGCA	CCCTCGCTTC	AGTACCTTAA	660
CAAGAAATAT	CCGACATAGG	AGAGGAGAAA	AGGTTGTGAT	CAATGTACCA	ATATTTAAGG	720
ACAAGAAATC	ACCATCTCCA	TTTATAGAAA	CAITTAATGA	GGATGATGAA	GCTTCAAGGG	780
CTTCTAAGCC	GGATCATATT	TACATGGATG	CCATGGGATT	TGGAATGGGC	AATTGCTGTC	840
TCCAGGTGAC	ATTCCAAGCC	TGCAATATAT	CTGAGGCCAG	ATACCTTTAT	GATCAGTTGG	900
CTACTATCTG	TCCAATTGTT	ATGGCTTTGA	GTGCTGCATC	TCCCTTTTAC	CGAGGCTATG	960
TGTGACAGAT	TGATTGTCCG	TGGGGAGTGA	TTTCTGCATC	TGTAGATGAT	AGAACTCGGG	1020
AGGAGCGAGG	ACTGGAGCCA	TTGAAGAAAC	ATACTATAG	GATCAGTAAA	TCCCGATATG	1080
ACTCAATAGA	CAGCATATTA	TCTAAGTGTG	GTGAGAAATA	TAATGACATC	GACTTGACGA	1140
TAGATAAAGA	GATCTACGAA	CAGCTGTGTC	AGGAAGGCAT	TGATCATCTC	CTGGCCGAGC	1200
ATGTTGTCTA	TCTCTTTATT	AGAGACCCAC	TGACACTGTT	TGAAGAGAAA	ATACACCTGG	1260
ATGATGCTTA	TGAGTCTGAG	CATTTTGAGA	ATATTGAGT	CACAAATTGG	CAGACAATGA	1320
GATTTAAGCC	CCCTCTCCCA	AACCTCAGCA	TTGATGGGAG	AGTAGAATTT	CGACCCATGG	1380
AGGTGCAATT	AACAGACTTT	GAGAACTCTG	CCTATGTGGT	GTTTGTGGTA	CTGCTACCCA	1440
GAGTGATCCT	TTCTTACAAA	TTGGATTGTC	TCATTCCACT	GTCAAAGGTT	GATGAGAAC	1500
TGAAGTAGAG	ACAGAAAAGA	GATGCTGTCT	TGCAGGGAAT	GTTTTATTTC	AGGAAAGATA	1560
TTTGCAAAAG	TGGCAATGCA	GTGGTGGATG	GTTGTGGCAA	GCCCCAGAAC	AGCAAGGAGC	1620
TCGCTCGAGA	GGAGTACACC	CTCATGAGCA	TAGACACCAT	CATCAATGGG	AAGGAAGGTG	1680
TGTTCTCGG	ACTGTACCCA	ATTCTGAACT	CTTACCTTGA	AAACATGGAA	GTGATGTGG	1740
ACACCATAGT	TAGTATTCTG	AACTACCTAA	AGCTAATTAA	GAAGAGAGCA	TCTGGAGAAC	1800
TAATGACAGT	TGCCAGATGG	ATGAGGGAGT	TTATCGCAAA	CCATCTTGAC	TACAGCAAG	1860
ACAGTGTCTA	AACGTATGAA	ATGAATTATA	GCCTTATTTT	GAAAGTGAAC	CAAAATTGCA	1920
ATGAATTATG	TGAATGCCCA	GAGTACTTGG	GATCAGCAAT	TAGGAAAGTA	AAATATAGTG	1980
GAAGTAAAC	TGACTCATCC	AACTAGACAT	TCTACAGAAA	GAAAAATGCA	TTATTGACGA	2040
ACTGGCTACA	GTACCATGCC	TCTCAGCCCG	TGTGTATAAT	ATGAAGACCA	AATGATAGAA	2100
CTGTACTGTT	TTCTGGGCCA	GTGAGCCAGA	AATGATTAA	GGCTTTCTTT	GSTAGGTAAA	2160
TCTAGAGTTT	ATACAGTGA	CATGTACATA	GTAAGTATT	TTGATTAAAC	AATGATTTTT	2220
AATAACATAT	CTAAAGTCAT	CATGAACTGG	CTTGATCAAT	TTTAAATTCT	TACTCTGGAG	2280
CAACCTACTG	TCTAAGCAGT	TTTGTAAATG	TACTGGTAAT	TGTACAATAC	TTGATTTCCA	2340
GAGTTAAAT	GTTTACTGTA	AATTTTGTG	CTTTTAAAGA	CTACCTGGGA	CCTGATTAT	2400
TGAAATTTTT	CTCTTTAAAA	ACATTTTCTC	TCGTTAATTT	TCCTTTGTCA	TTTCTTTGT	2460
TGTCTACATT	AAATCACTTG	AATCCATTGA	AAGTGCTTCA	AGGGTAATCT	TGGGTTTCTA	2520
GCACCTTATC	TATGATGTTT	CTTTTGCAAT	TGGATAATC	ACTTGGTCAC	CTTGCCCCAA	2580
GCTTTCCCTT	CTGAATAAAT	ACCCATTGAA	CTCTGAAAAA	AAAAA		

70

Seq ID NO: 294 Protein sequence:
Protein Accession #: NP_001489

75

1	11	21	31	41	51	
MGILLSQSPL	SWEETKRHAD	HVRRHIGILQF	LHIYHAVKDR	HKDVLRKWDG	VEYMLVSFDH	60
ENKKVRLVLS	GEKVELTLQE	KGERTNPNHP	TLWRPEYGSY	MEGTPEQGPY	GTMSEPNVTV	120
EAMNRKRRE	ATSLIENQA	LCTITSFPLR	GCPGFTLPEV	KPNPVEGGAS	KSLFFPDEAI	180
NKHFPRSTLT	RNIHRRGEK	VVINVPFKD	KNTPSPIET	PTEDEASRA	SKPDHIYMDA	240
MGFGMGNCCL	QVTPQACSI	EARLYLDQLA	TICPIVMAIS	AASPFYRGYV	SDIDCRWGI	300
SASVDRRTRE	ERGLEPLKNN	NYRISKSRVD	SIDSVLSKCG	EKYNDIDLTI	DKEIYQLLO	360
EGIDHLLAQH	VAHLFIRDPL	TLFEKIHLD	DANESDHFFN	IQSTNWTMR	FKPPPPNSDI	420
GWRFVFRPME	QLTDFENSA	YVVFVLLTR	VILSYKLDLF	IPLSKVDENM	KVAQKRDVIL	480
QGMFYFRKDI	CKGNAVVDG	CGKAQNSTEL	AAEYTLMSI	DTIINGKEGV	FPGLIPILNS	540
YLENMEVDVD	TRCSILNYLK	LIKKRASGEL	MTVARWNRFP	IANHPDYKQD	SVITDEMNYN	600
LILKCNQIAN	ELCECPBELLG	SAFRKVKYSG	SKTDSSN			

85

Seq ID NO: 295 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-816

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
      GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120
      CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
10     GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGG CCGTCTCTCC TCCAGCAAGG 240
      GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
      CCCAGGGAAT GTGACAGTCC TTGCTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
      AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
      ATGACAGGAC ATGCTATTCC ACCCAGCCAA TTGGATTCTC AGATTGATGA CTTCACTGGT 480
15     TTCAGCAAAG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
      ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAAGC 600
      CAACAGGAAA TTAATGCTGA TATAAAACGT AAATTAGTGA AGGAACTCCG ATGCGTTGGA 660
      CAAAAATATG AAAAAATCTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
      AAGCGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
20     AAGCACCTTA AGAAGAAACT GAAACGTATG ATTTGAGAAT ACTTGTCCCT GGAGGATTAT 840
      CACACCCCAA ATGCATAATC TCGTTAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
      TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCCTGGCAT ATGTTACCGA ATCAAAATAGC 960
      CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTT TTTTCCCAA AGCATTTTAT 1020
      TTGAAAGGAT AACTTGTGTT TTGGTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
25     TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A

```

Seq ID NO: 296 Protein sequence:
Protein Accession #: Eos sequence

```

30     1      11      21      31      41      51
      |      |      |      |      |      |
      MTDKTEKVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKEKKLMT 60
      GHAIPPSQLD SQIDDPFGPS KDRMMQKPGS NAPVGGNVTS SPSGDDLECR ETASSPKSQR 120
      EIDNADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIEAA RCMRRDFVKH 180
35     LKKKLKRLMI

```

Seq ID NO: 297 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-815

```

40     1      11      21      31      41      51
      |      |      |      |      |      |
      AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
      GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120
      CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
      GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGG CCGTCTCTCC TCCAGCAAGG 240
      GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
      CCCAGGGAAT GTGACAGTCC TTGCTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
      AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
      ATGACAGGAC ATGCTATTCC ACCCAGCCAA TTGGATTCTC AGATTGATGA CTTCACTGGT 480
      TTCAAGCAAG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
      ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAAGC 600
      CAACAGGAAA TTAATGCTGA TATAAAACGT AAATTAGTGA AGGAACTCCG ATGCGTTGGA 660
      CAAAAATATG AAAAAATCTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
      AAACGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
      AAGCACCTTA AGAAGAAACT GAAACGTATG ATTTGAGAAT ACTTGTCCCT GGAGGATTAT 840
      CACACCCCAA ATGCATAATC TCATTAAATG TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
      TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCCTGGCAT ATGTTACCGA ATCAACTGCG 960
      CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTT TTTTCCCAA AGCATTTTAT 1020
      TTGAAAGGAT AACTTGTGTT TTGGTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
      TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A

```

Seq ID NO: 298 Protein sequence:
Protein Accession #: Eos sequence

```

65     1      11      21      31      41      51
      |      |      |      |      |      |
      MTDKTEKVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKEKKLMT 60
      GHAIPPSQLD SQIDDPFGPS KDRMMQKPGS NAPVGGNVTS SPSGDDLECR ETASSPKSQQ 120
      EIDNADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIEAA RCMRRDFVKH 180
70     LKKKLKRLMI

```

Seq ID NO: 299 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-815

```

75     1      11      21      31      41      51
      |      |      |      |      |      |
      AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
      GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCCTCCCCA TGCACATCAC TCCCAGGTGC 120
      CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
      GGGAGGCGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGG CCGTCTCTCC TCCAGCAAGG 240
      GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
      CCCAGGGAAT GTGACAGTCC TTGCTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
      AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420
      TGACAGGACA TGCTATTCCA CCCAGCCAAAT TGGATTCTCA GATTGATGAC TTCCTGTTT 480
80     85

```


5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TCAGCAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540
CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAGGCC 600
AACCAAGAAAT TAATGCTGAT ATAAAACGTA AATTAGTGAA GGAACCTCGA TGCCTGGAC 660
AAAAATATGA AAAAATCTTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720
AAGCATTGTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTTA 780
AGCACCTTAA GAAGAAACTG AACGATATGA TTTGAGAATA CTGTCCCTCG GAGGATTATC 840
ACACCCCAAA TGCATAATCT CATTAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTTT 900
TCTACAATGG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTACCGAA TCAACTGGCC 960
TTCCAGAGGC TAAGAAATTT CTGTTAGTAA AAGATGTTCT TTTTCCCAAA GCGTTTATT 1020
TGAAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCAC CTGTGCTGGT AGATATTATT 1080
AACCCATTAG GTAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 300 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKAKKLMT 60
GHAIPPSQLD SQIDDPFTGFS KDRMMQKPGS NAPVGGNVT SFGSDDLLECR ETASSPKSQQ 120
EINADIKRKL VKELRCVGQK YEKIFEMLEB VQGPTAVRKR FFSIIKEAA RCMRRDFVKH 180
LKKLKRMI

Seq ID NO: 301 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-812

1 11 21 31 41 51
AGTGTTCGCG TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCTCTCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGACATCAC TCCAGGTGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGCGCC ACAACTTCAC TGCCATTGTT TGAGGTGCG CGTCTCTCC TCCAGCAAGG 240
GAACCAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAGCTTA 420
TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480
TCAGCAAGA TGGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAATGTTA 540
CCAGCAATTT CTCTGGAGAT GACCTAGAAT GCAGAGGAAT AGCCTCCTCT CCCAAAGGCC 600
AACCAAGAAAT TAATGCTGAT ATAAAATGTC AAGTAGTGAA GGAAATCOGA TGCCTGGAC 660
AATATGAAA AATCTTCGAA ATGCTTGAAG GAGTGCAAG ACCTACTGCA GTCAGGAAAC 720
GATTTTTTGA ATCCATCATC AAGGAAGCAG CAAGATGTAT GAGACGAGAC TTTGTTAAGC 780
ACCTTAAGAA GAAACTGAAA CGTATGATTT GAGAATACTT GTCCCTGGAG GATTATCACA 840
CCCCAATGC ATAACTCAT TAATGATTGA GGAGAGAAAA GGATCAGATT GCTGTTTTCT 900
ACAAATGGAG ATGATTTTGC TGAAGTCTCC TGGCATATGT TACCGAATCA ACTGGCCCTC 960
CAGAGGCTAA GAAATTTCTG TTAGTAAAG ATGTTCTTTT TCCCAAGCG TTTTATTGTA 1020
AAGGATAACT TGTGTTTGG TTATTTGTA TTCCACCTG TCGTGGTAGA TATTATTAA 1080
CCATTAGGTA AATACTATTA CAGTCGTGGT TTCTGCA

Seq ID NO: 302 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKEKLMT 60
GHAIPPSQLD SQIDDPFTGFS KDRMMQKPGS NAPVGGNVT SFGSDDLLECR GIASSPKSQQ 120
EINADIKQV VKELRCVGQY EKI FEMLEGV QGPTAVRKR FFSIIKEAAR CMRRDFVKHL 180
KKLKRMI

Seq ID NO: 303 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 247-815

1 11 21 31 41 51
AGTGTTCGCG TGGGACAGGC ACGCTGTGGC TGGCTACTTC CCTTCTCTCC ATCCCCCTTG 60
GGCCAAACAG GATCGGTGCT TCTGGTGAGA CGTCTCCCCA TGACATCAC TCCAGATGC 120
CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGCGCC ACAACTTCAC TGCCATTGTT TGAGGTGCG CGTCTCTCC TCCAGCAAGG 240
GAACCAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420
TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGGTT 480
TCAGCAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540
CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAGGCC 600
AACCAAGAAAT TAATGCTGAT ATAAAACGTA AATTAGTGAA GGAACCTCGA TGCCTGGAC 660
AAAAATATGA AAAAATCTTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720
AAGCATTGTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTTA 780
AGCACCTTAA GAAGAAACTG AACGATATGA TTTGAGAATA CTGTCCCTCG GAGGATTATC 840
ACACCCCAAA TGCATAATCT CGTTAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTTT 900
TCTACAATGG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTACCGAA TCAACTGGCC 960
TTCCAGAGGC TAAGAAATTT CTGTTAGTAA AAGATGTTCT TTTTCCCAAA GCGTTTATT 1020
TGAAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCAC CTGTGCTGGT AGATATTATT 1080
AACCCATTAG GTAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 304 Protein sequence:
Protein Accession #: Eos sequence

1 11 21 31 41 51
5 MTDKTEKIVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKAKKLMT 60
GHAIPPSQLD SQIDDFTFPS KDRMMQKPGS NAPVGGNVT SFGDDLECR ETASSPKSQ 120
EINADIKRKL VKELRCVGQK YEKIFEMLLEG VQGPTAVRKR FFESIIKEAA RCMRRDFVKH 180
LKKKLKRLMI

Seq ID NO: 305 DNA sequence
Nucleic Acid Accession #: Eos sequence
10 Coding sequence: 87-689

1 11 21 31 41 51
15 CGTGGAGGCA GCTAGCGCGA GGCTGGGGAG CGCTGAGCGG CGCGTGTGTG CCTGGGCTGC 60
CCAGACTAGC GAACAATACA GTCAGGATGG CTAAGGATGA CCCCAAGAAA CCAAGGGGCA 120
AGATGTCCCG TTATGCCTTC TTTGTGCAGA CATGCAGAGA AGAACATAAG AAGAAAAACC 180
CAGAGGTCCC TGTCAAATTT GCGGAATTTT CCAAGAAAGT CTCTGAGAGG TGGAAAGCGA 240
TGTCGGGGAA AGAGAAATCT AAATTTGATG AAATGGCAAA GGCAGATAAA GTGCGCTATG 300
20 ATCCGGAAAT GAAGGATTAT GGACCAGCTA AGGGAGGCAA GAAGAAGAAG GATCCTAATG 360
CTCCAAAAGG GCCACGCTCT GGATTCTTCC TGTTCTGTTC AGAATTCGCG CCCAAGATCA 420
AATCCACAAA CCCCGGCATC TCTATTGGAG ACGTGGCAAA AAAGCTGGGT GAGATGTGGA 480
ATAATTTAAA TGACAGTGAA AAGCAGCCTT ACATCACTAA GCGGCAAGG CTGAAGGAGA 540
AGTATGAGAA GGATGTGTCT GACTATAAGT CGAAAGGAAA GTTTGATGGT GCAAAGGGTC 600
CTGCTAAAGT TGCCCGGAAA AAGGTGGAAG AGGAAGATGA AGAAGAGGAG GAGGAAGAAG 660
25 AGGAGGAGGA GAGAGGAGGAG GATGAATAAA GAAACTGTCT ATCTGTCTCC TTGTGAATAC 720
TTAGAGTAGG GGAGCGCGGT AATTGACACA TCTCTTATTT GAGAAGTGTC TGTGCCCCTC 780
ATTAGTTTAA ATTACAAAT TTGATCACGA TCATATTGTA GTCTCTCAAA GTGCTCTAGA 840
AATTGTCACT GGTTTACATG AAGTGGCCAT GGGTGTCTGG AGCACCTCGA AACTGTATCA 900
AAGTTGTACA TATTTCCAAA CATTTTAAAA ATGAAAAGGC ACTCTGTGT TCTCTCACT 960
30 CTGTGCACTT TGCTGTGGT GTGACAAGGC ATTTAAAGAT GTTCTGGGCA TTTCTTTTT 1020
ATTTGTAAAG TGGTGGTAAC TATGGTTATT GGCTAGAAAT CCTGAGTTTT CAACGTGATA 1080
TATCTATAGT TTGTAAAAAG AACAAAACAA CCGAGACAAA CCTTGTATGC TCCTGTCTCG 1140
GCGTTGAGGC TGTGGGGAAG ATGCTTTTG GAGAGGCTG TAGCTCAGGG CGTGCACTGT 1200
GAGGCTGGAC CTGTTGACTC TGCAGGGGGC ATCCATTAG CTTCAGGTTG TCTTGTCTTC 1260
35 GTATATAGTG ACATAGCAAT CTGCTGCCAT CTTAGCTGTG GACAAAGGGG GGTGAGCTGG 1320
CATGAGAAAT TTTTTTTTTT TAAGTGCGGT AGTTTTTAAA CTGTTTGTGT TTAACAAAC 1380
TATAGAACTC TCTATTGTCA GCAAGCAAAA GAGTCACTGC ATCAATGAAA GTTCAAGAAC 1440
CTCCTGTACT TAAACACGAT TCGCAACGTT CTGTATTATTT TTTGTATGT TTAGAATGCT 1500
GAAATGTTTT TGAAGTTAAA TAAACAGTAT TACATTTTAA AAACCTCTCT CTATTATAAC 1560
40 AGTCAATTTT TGACTCACAG CAGTGAAACA ACCCCCACTC CATTGTATTT GGAGACTGGC 1620
CTCCCTATAA ATGTGGTAGC TTCTTTTATT ACTCAGTGGC CAGCTCACTT AGGGCTGAGA 1680
TGAAGGAGAG GGCTACTTGA AGCTACTGTG TGATTTTGT TGTGTCTGAG TGGCATTGAG 1740
ATGAAGTCTG GAGGAGTTAG GAGAACGACA TAGGCAAGGT TCAGCAGCCT TCCAAGGTAT 1800
AGGAAGGTGG GTGATTAGGA CTGAGGCTAT CTAGGTTTAA CTTTGTGCCC ACCTCCACCC 1860
45 CCTATTTTGT GGGGCCAAAT GCATTGCTAA ACAGCAATTT CAGAGTGTAT GGTGTGTCAA 1920
AAATTAAAGC CTTATTGTTT TTCTCTTTCA CCCCTACCCC CGGTGCTCCT GGCACATATC 1980
ACATTATTTG TGTGCCCCAA CATTTGGGTT CTGAGCCTG CTGCTGGTCT CCGGATGCC 2040
AGTGAGGGTA TGTGGGATGG GGTGGTGGGG TAGGGGACGG TATCCTTTTT TTGCTCTTAC 2100
50 TTGGAACAC CAAACACCCC AAGGAAGATG ATAGGCTCCA TCTTGGGCCA CCGTGGCTAT 2160
AGGGCAGGCT AATGGAATCA ACCATTTCTG AGCACTAAAT GTATCATGAA AAGTTGAATG 2220
GCCTGCTCAT AAGTTTAGCT CATTCACTGG AAATGTAGAT TGATGTTCAA TGTAAACTG 2280
GAAGGAGCTT GGTTTGTGTG TCAGTGGTTA TATTAGTGGG TAGTGTAACA TTTTATCCAG 2340
GTTGGGGTGA GGGGAGATGG CCACAGTAGC AAGTGGTGAC ACTAAATACC ATTTTGAAGG 2400
CTGATGTGTA TATACATCAT TACTGTCCGT AGCAATGAAG GATACAGTAC TGTGTTGTGG 2460
55 GTGAGTGGTA CTATTGCCCC GCAATTAATAT TTGGGTGTGT ATGTTTGAAG CTATGAAACA 2520
CGCAGGAGTG TTTTGTGTCT ATTAATTTTA AGAGAAAGCA GCTTTTCTCT AAAATTCAC 2580
GTTGAGAAAC TTGCATGTCT GGAGGCGGTG TCCTCTCCGC CCGTGGGT CCGGATGAG 2640
TACGAGTTAT GGTCAACGTC ACAGCCTGAT CTCTTATGTC TTCATAGCCA TTCGCTCTCC 2700
60 CATCAGAACT GTTTGTCTCT AATGTGTTC TCTAGTTCTA GAAATGACC ACTAATTTAA 2760
AAAACCTCGT TGTGAGGTTT GCCCAGAGGC ACTTGTTCCT GAAATTTCCC TCCTGCTTCA 2820
GCCATGTCCT TGTCACTTGG CATTTAAGC TAAAGCTTTA GCTTCCCAAT TCGTGATGTG 2880
CTAGGCCAAG ATTCGGGAGC TTTTCCAGC CTCGTCAAAT ATGGAAGAGA AACAACTGCG 2940
GGTCAAAAGG GAGTGATTG TTAAGTGGTG CGCGTCTATC TCATAACTAG ATGTACCAAC 3000
CAGGGAAGGG CCAAGGATGG AAAGGGGTAA CTTTGTGTCT TCCAAAGTAG CTAAGCAGAA 3060
65 GTGGGGGAGC AGTTTAGCCA GATGATCTTT GATTAGGCAA ACATTGAGTT TTAAGAGGC 3120
TGTCAGTTG AGGCCACTTG GTCCATTAGC TGGGGCAGCA AGATCACTAC TCAACGTTTT 3180
CACACTGTGG CAAGATTGCT CTTCTAGTGG AATAATGCC TAGTTTCTCT GAGATGATG 3240
AAGTGGCATG ATGTTACCTA AGGCTTAGGC TTAGCTTGAT TTCTGGGCCC ACTGCTGTG 3300
70 TTCTTAAGAT GCCAACCTGT TGCTTTTTTT TTTTCTTCC CCCATTAAA AGGATAGTAC 3360
CTACTCCCTC TAACCACTCT ACCCACTTCT TGAATGACAT TTTATCCTC GGAAAGAAC 3420
AGGCTGTGAT GTAGTACTA TTGTCTGTGT CTCCTGTGTG TGTCTGTCT TGTCAAAAT 3480
GTATTTGGGG ACGTTGGATG CATTCAATTT CTGTAATAAA G

Seq ID NO: 306 Protein sequence:
Protein Accession #: NP_005333.1

1 11 21 31 41 51
80 MAKGDPKPKK GKMSAYAFFV QTCREEHKKK NPEVPVNFAR FSKKCSERWK TMSGKEKSKF 60
DEMAKADKVR YDREMKDYGK AKGGKKKKDP NAPKRPPSGF FLFCSEPRPK IKSTNPGLSI 120
GDVAKKLGEN WNNLNDSEKQ PYITKAARKL ERYEKDVADY RSKGKFDGAK GPAKVARKKV 180
EEEEEEEEEE EEEEEEEEEE

Seq ID NO: 307 DNA sequence
Nucleic Acid Accession #: NM_022342
85 Coding sequence: 1..2178

	1	11	21	31	41	51	
	ATGGGTACTA	GGAAAAAAGT	TCATGCATTT	GTCGGTGTCA	AACCCACCGA	TGACTTTGCT	60
5	CATGAAATGA	TCAGATACGG	AGATGACAAA	AGAAGCATTG	ATATTCACCT	AAAAAAAGAC	120
	ATTGCGAGAG	GAGTTGTCAA	TAACCAACAG	ACAGACTGGT	CGTTTAAGTT	GGATGGAGTT	180
	TTCAAGATG	CCTCCAGGA	CTTGGTTTAT	GAGACAGTTG	CAAAGGATGT	GGTTTCTCAG	240
	CCCTCCGATG	GCTATAATGG	CACCATCATG	TGTTATGGGC	AGACGGGAGC	TGGCAAGACA	300
10	ACACCAATGA	TGGGGGCAAC	TGAGAAATAC	AAGCACCGGG	GGATCCTCCC	TGTCGCCCTG	360
	AGCAGGTTT	TTAGGATGAT	CGAAGAACGC	CCCACACATG	CCATCACTGT	GGGTGTTTCC	420
	ACTTGGAAA	TCTATAATGA	GAGCCTGTTT	GATCTCCTGT	CCACTCTGCC	CTATGTTGGA	480
	CCTCAGTCA	CACCAATGAC	CATCGTGGAA	AACCTCAAG	GAGTCTTCAT	TAAGGGCTTG	540
	CAGTTACC	TCACAAGTCA	GGAGGAGGAT	GCATTACGCC	TCCTTTTGA	GGGTGAGACC	600
15	ACAGGATTA	TAGCCTCCCA	CACATGAAC	AAAACTCTT	CCAGATCACA	CTGCATTTTC	660
	CCATCTACT	TAGAGGCCCA	TTCCCGGACC	TTATCAGAGG	AAAAGTACAT	CACCTCCAAA	720
	TTAACTTGG	TGGATCTGGC	AGGCTCAGAG	AGGCTGGGGA	AGTCTGGGTC	TGAGGGCCAA	780
	TCCTGAAGG	AAGCCACCTA	CATCAACAAA	TCGCTCTCAT	TCCTGGAGCA	GGCCATCATT	840
	CCCTTGGGG	ACCAGAAGCG	GGACCAATC	CCCTTTCGGC	AGTGAAGCT	CACCCAOGCT	900
20	TGAAGGACT	CGTTAGGGGG	AAACTGCAAT	ATGCTCCTCG	TGACAAACAT	CTATGGAGAA	960
	CTGCCCACT	TAGAAGAAAC	GCTATCTTCA	CTGAGATTTG	CCAGCAGGAT	GAAGCTAGTC	1020
	CCACTGAGC	CTGCCATCAA	TGAAAAGTAT	GATGCTGAGA	GAATGGTCAA	GAACCTGGAG	1080
	AGGAAGTAT	CACCAATGAC	GCAGGAGCTG	GCTATCCATG	ACAGCCTGAC	CAACCGCACC	1140
	TTGTGACCT	ATGACCCCAT	GGATGAAATC	CAGATTGCTG	AGATCAACTC	CCAGGTGCGG	1200
25	GGTACCTGG	AGGGGACATC	GGACGAGATC	GACATAATCA	GCCTTAGACA	GATCAAGGAG	1260
	TGTTCAACC	AGTTCCGGGT	GGTTCTGAGC	CAACAGGAAC	AGGAAGTGA	TGCCACTTTG	1320
	GCAGGAAGT	ACACCCCTAT	TGACAGGAAT	GACTTTGCG	CCATTTCTGC	TATCCAGAA	1380
	CGGGGCTTG	TGGATGTTGA	TGGCCACCTA	GTGGGTGAGC	CTGAAGGACA	AAACTTTGGA	1440
	TCGGAGTGG	CCCTTTCTC	TACCAAACT	GGGAAGAAAG	CAAAGTCAA	GAAGACATTC	1500
30	AAGAGCCAC	TCAGGCCCGA	CACCCCAACC	TCCAAACAG	TGGCCTTTGA	GGAGTTTAA	1560
	ATGAGCAAG	GTAGTGAGAT	CAACCGAATT	TTCAAGAAA	ACAATCCAT	CTTGAATGAA	1620
	GGAGGAAAA	GGGCGACGGA	GACCAACAG	CACATCAATG	CCATCAAGCG	GGAGATTGAT	1680
	TGACCAAGG	AGGCCCTGAA	TTTCCAGAA	TCACTACGGG	AGAAGCAAGG	CAAGTACGAA	1740
	ACAAGGGGC	TGATGATCAT	CGATGAGGAA	GAATTCCTGC	TGATCCTCAA	GCTCAAAGAC	1800
35	TCAAGAAAG	AGTACCGCAG	CGAGTACCAG	GACCTGCGTG	ACCTCAGGGC	TGAGATCCAG	1860
	ATTGCCAGC	ACCTAGTGGA	TCAGTGTGCG	CACCGCCTGC	TCATGGAATT	TGACATCTGG	1920
	ACAATGAGT	CCTTTGTGAT	CCCTGAGGAC	ATGCAGATGG	CACTGAAGCC	AGGCGGCAGC	1980
	TCGGGCCAG	GCATGGTCCC	TGTGAACAGG	ATTGTGCTC	TGGGAGAAGA	TGACCAAGAC	2040
	AATTACAGC	AGCTGCAGCA	GAGGTGCTT	CCTGAGGGCC	CTGATTCAT	CTCCTTCTAC	2100
40	ATGCCAAAG	TCAAGATAGA	GCAGAAGCAT	AATTACTTGA	AAACCATGAT	GGGCTCCAG	2160
	AGGCACATA	GAATAATAG					

Seq ID NO: 308 Protein sequence:
Protein Accession #: NP_071737

	1	11	21	31	41	51	
	MGTRKKVHAP	VRVKPTDDFA	HEMIRYGGDK	RSIDHLLKDD	IRRGVVNNQQ	TDWSFKLDGV	60
	LHDASQDLVY	ETVAKDVVSG	ALDGYNGTIM	CYQGTGAGKT	YTMGATENY	KHRGILPRAL	120
50	QQVFRMIEER	PTHATVVRVS	YLEIYNESLF	DLLSTLPYVG	PSVTPMTIVE	NPQGVFIKGL	180
	SVHLTSQEEED	AFSLTFEGET	NRIIASHTMN	KNSSRSCHIF	TIYLEAHSRT	LSEKYYITSK	240
	INLVDLAGEB	RLGKSGSEGG	VLKEATYINK	SLSFLEQAI	ALGDQKRDHI	PFRQCKLTHA	300
	LKDSLGNCN	MVLVTNIYGE	AAQLEETLSS	LRFAARMKLV	TTEPAINEKY	DAERMVKNLE	360
	KELALLKQEL	AIHDSLTNRT	FVTYDPMDEI	QIAEINSQVR	RYLESTLDEI	DIISLRQIKE	420
	VFNQPRVLS	QQQEVESTL	RRKYTLIDRN	DFAAISAIQK	AGLVDDVGH	VGEPEGQNF	480
55	LGVAFFSTKF	GKKAKSKKTF	KEPLRPDTTP	SKPVAFEFPK	NEQGSSEINR	PKENKSIKLE	540
	RRKRASSETQ	HINAIKREID	VTKALNFOK	SLREKQGYE	NKGLMIDEB	EPLILKLKD	600
	LKKQYRSEYQ	DLRLDLRAIQ	YQHLVDQCR	HRLIMEFDIW	YNESFVIPED	MQMALRPGGS	660
	IRPGMVFVNR	IVSLGEDDQD	KFSQLQQRVL	PEGPDSISPY	NAKVKIEQKH	NYLKTMMGLQ	720
60	QAHRK						

Seq ID NO: 309 DNA sequence
Nucleic Acid Accession #: CAT cluster

	1	11	21	31	41	51	
	TTTTTTTTTT	TTTTTTTAA	TGCTGTCTGT	CATGCTCTGT	CTACCAAGGT	GAATTCCAA	60
	AAATTTCTGC	ATAGCAATTT	TAGCCAAAAC	TATATATGTT	CTGGGGAGGA	TAGGCATAGG	120
	CACATTGAAG	ACCAAAGGAA	AGAGTGAAGA	AGTGTAGTTG	GGTCATTGTG	AATGGATGTT	180
70	TAGATTGTCA	AGAAAAGTGG	GCCAGAGGCC	CCACCTCACA	CTAGGACGGC	AATTGCCCTCT	240
	CATTAGTATC	TCAGGCACCA	TGGGTCTTAT	TTGGTGTGAT	AAGAAACACC	CTCAACAAAG	300
	TAATGAACCC	TCAGCCTCCA	GCTTCTCTTC	TTCCGGGATTC	TTCTTAGGGC	CTCCTTTTTC	360
	CTTTTATGTT	TCCAGTACCC	TGAATTTCTT	ATTCCCATCC	CCATTAAAA	TCTGCTTCAA	420
	AGAAAAAACA	AGAAGGACAC	ATTCACTTTA	AGATCCAAAT	GAATGATAAG	AGCTTAAAAAC	480
75	ATTATACCTA	TCAGTATTAT	TTGCATTTTT	ATAGAAACCA	AAACCATATT	TCAACAAC	

Seq ID NO: 310 DNA sequence
Nucleic Acid Accession #: NM_018622.2
Coding sequence: 1-1140

	1	11	21	31	41	51	
	ATGGCGTGGC	GAGGCTGGGC	GCAGAGAGGC	TGGGGCTGCG	GCCAGGCGTG	GGGTGCGTGG	60
	GTGGGGCGGC	GCAGCTCGGA	GGAGCTCACT	GGGTCTCTAA	CCCCGCGGCA	GCTCCTCGGA	120
	CGCAGGTTTA	ACTTCTTTAT	TCAACRAAAA	TGCGGATTCA	GAAAAGCACC	CAGGAAGGTT	180
85	GAACCTCGAA	GATCAGACCC	AGGGACAAGT	GGTGAAGCAT	ACAAGAGAAG	TGCTTTGATT	240
	CCTCCTGTGG	AAGAAACAGT	CTTTTATCCT	TCTCCCTATC	CTATAAGGAG	TCTCATAAAA	300
	CCTTTATTTT	TTACTGTTGG	GTTTACAGGC	TGTGCATTGG	GATCAGCTGC	TATTTGGCAA	360

	TATGAATCAC	TGAAATCCAG	GGTCCAGAGT	TATTTTGATG	GTATAAAAGC	TGATTGGTTG	420
	GATAGCATAA	GACCACAAAA	AGAAGGAGAC	TTCAGAAAGG	AGATTAAACAA	GTGGTGGAAAT	480
	AACCTAAGTG	ATGGCCACGG	GACTGTGACA	GGTATTATAG	CTGCAAAATGT	CCTTGTATTTC	540
5	TGTTTATGGA	GAGTACCTTC	TCTGCAGCGG	ACAATGATCA	GATATTTCAC	ATCGAATCCA	600
	GCCTCAAAGG	TCCTTTGTTC	TCCAATGTG	CTGTCAACAT	TCAGTCACCT	CTCCTTATTT	660
	CACATGGCAG	CAAAATATGT	TGTTTTGTGG	AGCTTCTCTT	CCAGCATAGT	GAACATTCTG	720
	GGTCAAGAGC	AGTTTCATGGC	AGTGTACCTA	TCTGCAGGTG	TTATTTCCAA	TTTGTGTCAGT	780
	TACCTGGGTA	AAGTTGCCAC	AGGAAGATAT	GGACCATCAC	TTGGTGCATC	TGGTGCCATC	840
10	ATGACAGTCC	TCGCAGCTGT	CTGCACTAAG	ATCCCAAGAG	GGAGGCTTGC	CATTATTTTC	900
	CTTCCGATGT	TCACGTTTAC	AGCAGGGAAT	GGCTTGAAG	CCATTATCSC	CATGGATACA	960
	GCAGGAATGA	TCCTGGGATG	GAATTTTTTT	GATCATGCGG	CACATCTTGG	GGGAGCTCTT	1020
	TTTGAATAT	GGTATGTTAC	TTACGGTCAT	GAACGTATTT	GGAAGAACAG	GGAGCCGCTA	1080
	GTGAAATCT	GGCATGAAAT	AAGGACTAAT	GGCCCCAAAA	AAGGAGGTGG	CTCTAAGTAA	

15

Seq ID NO: 311 Protein sequence:
Protein Accession #: NP_061092.2

20	1	11	21	31	41	51	
	MAWRGMAQRG	WGCGQAWGAS	VGGRSCEELT	AVLTTPQLLG	RRFNFFIQOK	CGFRKAPRKV	60
	EPFRSDPGTS	GEAYKRSALI	PFVEETVFYP	SPYPISRLIK	PLFFTVGPTG	CAFSGAAIWQ	120
	YESLKSRSVQS	YFDIKADWL	DSIRPQKEGD	FRKEINKWVN	NLSDGQRTVT	GIIAANVLVF	180
	CLMRVPSLQR	TMIRYFTSNP	ASKVLCSPML	LSTFHSFSLP	HMAANMYVLW	SPSSSIVNIL	240
25	GQEQFMAVYL	SAGVISNFPV	YLKQVATGRY	GPSLGAAGAI	MTVLAAVCTK	IPBGRALIIIP	300
	LPMPFTPTAGN	ALKAIIMADT	AGMILGWKFP	DHAHLGGAL	PGIWTYVTYH	ELIWNKREPL	360
	VKIWHIEIRTN	GPKKGGGSK					

30

Seq ID NO: 312 DNA sequence
Nucleic Acid Accession #: NM_000625
Coding sequence: 195..3656

35	1	11	21	31	41	51	
	CTCTCGGCCA	CCTTTGATGA	GGGGACTGGG	CAGTCTCTAGA	CAGTCCCGAA	GTCTCTCAAGG	60
	CACAGGTCTC	TTCTCTGGTTT	GACTGTCTCT	ACCCCGGGGA	GGCAGTGCAG	CCAGCTGCAG	120
	GCCCCACAGT	GAAGAACATC	TGAGCTCAAA	TCCAGATAAG	TGACATAAGT	GACCTGCTTT	180
	GTAAGCCCAT	AGAGATGGCC	TGTCTTGTGA	AATTTCTGTT	CAAGACCAAA	TTCCACCAGT	240
40	ATGCAATGAA	TGGGGAAGAA	GGCATCAACA	ACAATGTGGA	GAAAGCCCCC	TGTGCCACCT	300
	CCAGTCCAGT	GACACAGGAT	GACCTTCAGT	ATCAACAACCT	CAGCAAGCAG	CAGAATGAGT	360
	CCCGCAGGCC	CCTCGTGGAG	ACGGGAAAGA	AGTCTCCAGA	ATCTCTGGTC	AAGCTGGATG	420
	CAACCCCAT	GTCTCTCCCA	CGGCATGTGA	GGATCAAAAA	CTGGGGCAGC	GGGATGACTT	480
	TCCAAGACAC	ACTTCAACAT	AAGGCCAAAG	GGATTTTAAC	TTGCAGGTCC	AAATCTTGCC	540
45	TGGGGTCCAT	TATGACTCCC	AAAAGTTTGA	CCAGAGGACC	CAGGGACAAG	CCTACCCCTC	600
	CAGATGAGCT	TCTACCTCAA	GCTATCGAAT	TTGTCAACCA	ATATTACGGC	TCCTCAAAG	660
	AGGCAGAAAT	AGAGAAACAT	CTGGCCAGGG	TGGAGCGGT	AACAAAGGAG	ATAGAAACAA	720
	CAGTAACCTA	CCAACCTAGC	GGAGATGAGC	TCATCTTCGC	CACCAAGCAG	GCCTGGCCCA	780
	ATGCCCCCAG	CTGCATTGGG	AGGATCCAGT	GGTCCAACT	GCAGGTCTTC	GATGCCCGCA	840
50	GCTGTTCCAC	TGCCCGGGAA	ATGTTTGAAC	ACATCTGCAG	ACACGTGCGT	TACTCCACCA	900
	ACAATGGCAA	CATCAGGTGC	GCCATCACCG	TGTTCCCCCA	GCGGAGTGAT	GGCAAGCAAG	960
	ACTTCCAGGT	TGGAATGCT	CAGCTCATCC	GCTATGCTGG	CTACAGATG	CCAGATGGCA	1020
	GCATCAGAGG	GGACCTTGCC	AACGTGGAAT	TCACTCAGCT	GTGCATCGAC	CTGGGCTGGA	1080
	AGCCCCAAGT	CGGCCGCTTC	GATGTGCTCC	CCCTGGTCTT	GCAGGCCAAT	GGCCGTGACC	1140
55	CTGAGCTCTT	CTGAAATCCA	CCTGACCTTG	TGCTTGAGGT	GGCCATGGAA	CATCCCAAT	1200
	ACGAGTGGTT	TCGGGAACTG	GAGCTAAAGT	GGTAGCCCTT	GCCTGCAGTG	GGCAACATGC	1260
	TGCTTGAGGT	GGGCGGCTGT	GAGTTCCAG	GGTGCCTCTT	CAATGGCTGG	TACATGGGCA	1320
	CAGAGATCGG	AGTCCGGGAC	TTCTGTGATG	TCCAGCGCTA	CAACATCTTG	GAGGAAGTGG	1380
	GCAGGAGAA	GGGCTCGGAA	ACGCACAAGC	TGGCTCTGCT	CTGGAAAGAC	CAGGCTGTGG	1440
60	TTGAGATCAA	CTTCTGTGTG	CTCCATAGTT	TCCAGAAAGC	GAATGTGACC	ATCATGGACC	1500
	ACCACTCGGC	TGCAGAAATC	TTTATGAAGT	ACATGCAGAA	TGAATACCGG	TCCCGTGGGG	1560
	GCTGCCCGGC	AGACTTGAAT	TGGCTGTGCT	CTCCCATGTC	TGGGAGCATC	ACCCCGGTGT	1620
	TTCAACAGGA	GATGCTGAAC	TACGTCTCTG	CCCCTTTCTA	CTACTATCAG	GTAGAGGCTT	1680
	GGAAACCCCA	TGCTGCGCAG	GACGAGAAGC	GGAGACCCAA	GAGAAGAGAG	ATTCCATTGA	1740
65	AAGTCTTGCT	CAAGCTGTGT	CTCTTTGCTT	GTATGCTGAT	GGCAGAGACA	ATGGCGTCCC	1800
	GAGTCAGAGT	CACCATCTCT	TTTGGCAGAG	AGACAGGAAA	ATCAGAGGCG	CTGGCTGGGG	1860
	ACCTGGGGGC	CTTATTTCAG	TGTGCCCTTC	ACCCCAAGGT	TGTCTGCATG	GATAAGTACA	1920
	GGCTGAGCTG	CCTGGAGGAG	GAAAGGCTGC	TGTTGGTGGT	GACCAAGTAC	TTTGGCAATG	1980
	GAGACTGCCC	TGGCAATGGA	GAGAACTGTA	AGAAATCGCT	CTTCACTGCT	AAAGAGCTCA	2040
70	ACAAACAAAT	CAGGTACGCT	GTGTTTGGCC	TGGGCTCCAG	CATGTACCTT	CGGTTCTGGG	2100
	CCTTTGCTCA	TGACATTTGAT	CAGAAGCTGT	CCACCTGGGG	GGCCTCTCAG	CTCACCCCGA	2160
	TGGGAGAAGG	GGATGAGCTC	AGTGGGCAGG	AGGACGCTTT	CCGAGCTGCG	GGCCTGCAGG	2220
	CCTTCAAGGC	AGCCTGTGAG	ACGTTTGATG	TCCGAGGCAA	ACAGCACATT	CAGATCCCCA	2280
	AGCTCTACAC	CTCCAAATGT	ACCTGGGACC	CGCACCACTA	CAGGCTCGTG	CAGGACTCAC	2340
75	AGCCTTTGGA	CCTCAGCAAA	GCCTCAGCA	GCATGCATGC	CAAGAACGCTG	TTCAACATGA	2400
	GGCTCAAAAT	TGGCAGAAAT	CTACAAAGTC	CGATATCCAG	CGGTGCCACC	ATCCTGGTGG	2460
	AACTCTCTCT	TGAGGATGGC	CAAGGCTTGA	ACTACCTGCC	GGGGGAGCAC	CTTGGGGTTT	2520
	GCCGAGGCAA	CCAGCCGGCC	CTGGTCCAAG	GCATCTTGGG	GCGAGTGGTG	GATGGCCCCA	2580
	CACCCACCA	GGCAGTGGCC	CTGGAGGCCC	TGGATGAGAG	TGGCAGCTAC	TGGGTGAGTG	2640
80	ACAGAGGCTC	GGCCCCCTGC	TCACTCAGCC	AGGCCCCCTC	CTACTTCTCT	GACATCACCA	2700
	CACCCCAAC	CCAGCTGCTG	CTCCAAAGC	TGGCCCAAGT	GGCCACAGAA	GAGCCTGAGA	2760
	GACAGAGGCT	GGAGGCCCTG	TGCCAGCCCT	CAGAGTACAG	CAAGTGAAG	TTCAACCAAC	2820
	GCCCCACATT	CCTGGAGGCT	CTAGAGGAGT	TCCCGTCCCT	GCGGCTGTCT	GCTGGCTTCC	2880
	TGCTTTCCCA	GCTCCCATCT	CTGAAGCCCA	GGTTCTACTC	CATCAGCTCC	CCCCGGGATC	2940
85	ACAGGCCCA	GGAGATCCAC	CTGACTGTGG	CCGTGGTTC	CTACCAACCC	CGAGATGGCC	3000
	AGGGTCCCC	GCACACCGCC	GTCTGCAGCA	CATGGCTCAA	CAGCTGAAG	CCCCAAGACC	3060
	CAGTGCCTTG	CTTTGTGCGG	AATGCCAGCG	GCTTCCACCT	CCCCGAGGAT	CCCTCCCATC	3120

CTTGCATCCT CATCGGGCCT GGCACAGGCA TCGCGCCCTT CCGCAGTTTC TGGCAGCAAC 3180
 GGCTCCATGA CTCCAGCAGC AAGGGAGTGC GGGGAGGCCG CATGACCTTG GTGTTTGGGT 3240
 CGCGCGCCCC AGATGAGGAC CACATCTACC AGGAGGAGAT GCTGGAGATG GCCCAGAAAG 3300
 GGGTGTCTGA TGGGGTGCAC ACAGCCTATT CCGCCTGCCC TGGCAAGCCC AAGGTCTATG 3360
 5 TTCAGGACAT CCTCGCGCAG CAGCTGGCCA GCGAGGTGCT CCGTGTGCTC CACAAGGAGC 3420
 CAGGCCACCT CTATGTTTGC GGGGATGTGC GCATGGCCCC GGACGTGGCC CACACCCCTGA 3480
 AGCAGCTGTT GGCTGCCAAG CTGAAATTGA ATGAGGAGCA GGTGAGGAC TATTCTTTTC 3540
 AGCTCAAGAG CCAAGAGCGC TATCAAGAG ATATCTTTGG TGCTGTATT CTCTACGAGG 3600
 10 CGAAGAAGGA CAGGGTGGCG GTGCAGCCCA GCAGCCTGGA GATGTGAGG CTCTGAGGGC 3660
 CTACAGGAGG GGTAAAGCT GCCGGCACAG AACTTAAGGA TGGAGCCAGC TCTGCATTAT 3720
 CTGAGGTAC AGGGCTGGG GAGATGGAGG AAGGTGATAT CCCCCAGCCT CAAGTCTTAT 3780
 TTCTCAAGC TTGCTCCCCA TCAAGCCCTT TACTTGACCT CTTAACAAGT AGCACCCCTG 3840
 ATTGATCGGA GCCTC

15 Seq ID NO: 313 Protein sequence:
 Protein Accession #: NP_000616

20 1 11 21 31 41 51
 MACPWKFLFK TKFHQYAMNG EKGINNVEK APCATSSPVT QDDLQYHNLS KQONESPQPL 60
 VETGKSPES LVKLDATPLS SPRHVRKNW GSGMTFQDTL HHKAGILTC RSKSLGSLIM 120
 TPKSLTRGPR DKPTPPDELL POAIEFVNQY YGSLKEAKIE EHLARVEAVT KEIETTUTYQ 180
 25 LTGDELIFAT KQAWRNAPRC IGRQWSNLQ VFDARSCTA REMPEHICRH VRYSTNNNGNI 240
 RSATITVPPQR SDGKHDFRWN NAQLIRYAGY QMPDGSIRGD PANVEFTQLC IDLGWPKKYG 300
 RFDVVLVLQ ANGRDPELFE IPPDLVLEVA MEHPKYEWFR EELKWKYALP AVANMLLEVQ 360
 GLEFPQCFN GWMYQTEIGV RDPFDVQRYN ILEEVGRRMG LETHKLASLW KDQAVVEINI 420
 AVLHSPQKQV VTIMDHSAE RSPMKYQNE YRSRGCPAD WTLVPPMSG SITPVPHQEM 480
 30 LNYVLSFFYY YQVEAWKTHV WQDEKRRPKR REIPLKVLVK AVLFACMLMR KTMASRVRT 540
 ILFATETGKS EALAWDLGAL FSCAPNPKV CNDKYRLSCL EERLLLVVT STFGNGDCPG 600
 NGELKKSFLP MLKELNNKFR YAVFGLGSSM YRPFCAFD IDQKLSHLGA SQLTFMGEGD 660
 ELSGQEDAFR SWAVQTFKAA CETFDVRGKQ HIQIPKLYTS NVTWDPHYR LVQDSQPLDL 720
 SKALSSMHAK NVFTMLRKS QNLQSPSSR ATILVELSCE DQGLNVLPG EHLGVCPGNQ 780
 35 PALVQGILER VVDGPTPHQA VRLEALDESG SYWVSDKRLP PCSLSQALTY FLDITTPPTQ 840
 LLLQLAQVA TEPEFRQRIE ALQCPSEYSK WKFTNSPTFL EVLEEFPSLR VSAGFLLSQL 900
 PILKPRFYSI SSPRDHTPTE IHLTVAVVTY HTRDQGGPLH HGVCSTWLSN LKQDPVPCF 960
 VRNAGFHLP EDPGHPCLIL GPGTGIAFPR SFWQQRLEDH QHKGVGRGRM TLVFGCRRPD 1020
 EDHIYQZML EMAQKGVLEA VHTAYSRLPG KPKVYVQDIL RQQLASEVLR VLEKEPHGLY 1080
 40 VCGDVRMARD VAHTLRQLVA AKLKLNEEQV EDYFFQLKSG KRYHEDIFGA VFPYEAKKDR 1140
 VAVQSSLEM SAL

45 Seq ID NO: 314 DNA sequence
 Nucleic Acid Accession #: XM_087254
 Coding sequence: 47..2332

50 1 11 21 31 41 51
 AGAGTACGTG TTTACAGATA AAACCTGGTAC ACTGACAGAA AATGAGATGC AGTTTCGGGA 60
 ATGTTCAAAT AATGGCATGA AATACCAAGA AATTAATGGT AGACTTGTAC CGAAGGACC 120
 AACACACAGC TCCTCAGAGG GAAACTTATC TTATCTTAGT AGTTTATCCC ATCTTAACAA 180
 CTTATCCCAT CTTACAACCA GTTCTCTTT CAGAACCACT CCTGAAATG AACTGAACT 240
 AATTAAAGAA CATGATCTCT TCTTTAAAGC AGTCAGTCTC TGTCACACTG TACAGATTAG 300
 55 CAATGTTCAA ACTGACTGCA CTGGTGATGG TCCCTGGCAA TCCAACCTGG CACCATGGCA 360
 GTTGGAGTAC TGTGATCTT CACCAGATGA AAAGGCTCTA GTAGAAGCTG CTGCAAGGAT 420
 TGTGATTGTG TTTATTGGCA ATTCTGAAGA AACTATGGAG GTTAAACTC TTGAAAACCT 480
 GGAACGGTAC AAACCTGCTC ATATTCTGGA ATTTGATTCA GATCGTAGGA GAATGAGTGT 540
 AATTGTTTCA GCACCTTCAG GTGAGAAGTT ATTATTGTCT AAAGGAGCTG AGTCATCAAT 600
 60 TCTCCCTAAA TGTATAGGTG GAGAAATAGA AAAAACCAGA ATTCATGTAG ATGAATTTCG 660
 TTTGAAAGGG CTAAGAATCT TGTGTATAGC ATATAGAAAA TTTACATCAA AAGAGTATGA 720
 GGAATAGAT AAACGCATAT TTGAAGCCAG GACTGCCTTG CAGCAGCGGG AAGAGAAAT 780
 GGCAGCTGTT TTCCAGTTCA TAGAGAAAGA CCTGATATTA CTGGAGCCA CAGCAGTAGA 840
 AGACAGACTA CAAGATAAAG TTCGAGAAAC TATTGAAGCA TTGAGAATGG CTGGATCAA 900
 65 AGTATGGGTA CTACTGGGG ATAAACATGA AACAGCTGTT AGTGTGAGTT TATCATGTGG 960
 CCATTTTCAT AGAACCATGA ACATCCTTGA ACTTATAAAC CAGAAATCAG ACAGCGAGTG 1020
 TGCTGAACAA TTGAGGCAGC TTGCCAGAAG AATTACAGAG GATCATGTGA TTCAGCATGG 1080
 GCTGGTAGTG GATGGGACCA GCCTATCTCT TGCCTCAGG GAGCATGAAA AACTATTAT 1140
 GGAAGTTTGC AGAAATTGTT CAGCTGTATT ATGCTGTCGT ATGGCTCCAC TGCAGAAAGC 1200
 70 AAAAGTAATA AGACTAATAA AAATATCACC TGAGAAACCT ATAACATTGG CTGTTGGTGA 1260
 TGGTGCTAAT GACGTAAGCA TGATAAAGA AGCCCATGTT GGCAATAGAA TCATGGGTAA 1320
 AGAAGGAAGA CAGGCTGCAA GAAACAGTGA CTATGCAATA GCCAGATTGA AGTTCCTCTC 1380
 CAAATTGCTT TTTGTTTCATG GTCATTTTAT TTATATTAGA ATAGCTACCC TTGTACAGTA 1440
 TTTTITTTAT AAGAATGTGT GCTTTATCAC ACCCCAGTTT TTATATCAGT TCTACTGTTT 1500
 75 GTTTTCTCAG CAAACATTGT ATGACAGCGT GTACCTGACT TTATACAATA TTTGTTTAC 1560
 TTCCCTACCT ATTCTGATAT ATAGTCTTTT GGAACAGCAT GTAGACCCCTC ATGTGTTACA 1620
 AAATAAGCCC ACCCTTTATC GAGACATTAG TAAAAACCGC CTCCTTAAGTA TTAACCAATT 1680
 TCITTTATTGG ACCATCCTGG GCTTCAGTCA TGCCITTTAT TTTCTTTTTC GATCCTATTT 1740
 ACTAATAGGG AAGATACAT CTCTGCTTGG AAATGGGCCA ATGTTTGGAA ACTGGACATT 1800
 80 TGGCATTGG GTCTTCACAG TCATGGTTAT TACAGTCACA GTAAAGATGG CTCTGGAAAC 1860
 TCATTTTGG ACTGTGATGA ACCATCTCGT TACCTGGGGA TCTATTATAT TTTATTTTGT 1920
 ATTTTCTCTG TTTTATGGAG GGATTCTCTG GCCATTTTTC GGCTCCGAGA ATATGTATTT 1980
 TGTGTTTAT CAGCTCCTGT CAAGTGGTTC TGCTTGGTTT GCCATAATCC TCATGGTTGT 2040
 TACATGTCTA TTTCTTGATA TCATAAGAA GGTCTTTGAC GCACACCTCC ACCCTACAAG 2100
 85 TACTGAAAAG GCACAGCTTA CTGAAACAAA TGCAAGTATC AAGTCTTGG ACTCCATGTG 2160
 CTGTTTCCCG GAAGGAGAG CAGCGTGTGC ATCTGTTGGA AGAATGCTGG AACGAGTTAT 2220
 AGGAAGATGT AGTCCAACCC ACATCAGCAG ATCATGGAGT GCATCGGATC CTTTCTATAC 2280
 CAACGACAGG AGCATCTTGA CTCTCTCCAC AATGAGTCA TCTACTGTTT AAAGGGGCAG 2340

	TAGTACTTTG	TGGGAGCCAG	TTACCTCTCT	TTCTCTAAAT	TCAGTGTGAT	CACCTGTGTA	2400
	ATGGCCACAC	TAGCTCTGAA	ATTAATTTCC	AAATCTTTG	TAGTAGTTCA	TACCCACTCA	2460
	GAGTTATAAT	GGCAAAACAA	CAGAAAGCAT	TAGTACAAGC	CCCTCCCAAC	ACCCCTTAAT	2520
5	TGAATCTGAA	CAATGTTAAA	TTTGAGAATA	AAGAGACATT	TTTCATCTCT	TGTGCTGTT	2580
	TGTCCTCTGT	GCTTATGGGA	CTCCTAATGG	CAITTCAGTC	TGTTGCTGAG	GCCATTATAT	2640
	TTTAATATAA	ATGTAGAAAA	AAGAGAGAAA	TCTTAGTAAA	GAGTATTTTT	TAGTATTAGC	2700
	TTGATTATTG	ACTCTCTTAT	TTAAATCTGC	TTCTGTAAAT	TATGCTGAAA	GTTTGCCCTG	2760
	AGAACTCTAT	TTTTTTTATTA	GAGTTATATT	TAAAGCTTTT	CATGGGAAAA	GTTAATGTGA	2820
10	ATACTGAGGA	ATTTTGGTCC	CTCAGTGACC	TGTGTTGTTA	ATTCAATTAAT	GCACTCTGAG	2880
	TTACAGAGAG	AAATTAGGAG	AATCAITTC	AACCAITTAAT	TACTGCAGTA	TGGGGAGTAA	2940
	ATTTATACCA	ATTCCTCTAA	CTGTACTGTA	ACACAGCCTG	TAAAGTTAGC	CATATAAATG	3000
	CAAGGGTATA	TCATATATAC	AAATCAGGAA	TCAGGTCOCT	TCACCGAACT	TCAAATTGAT	3060
	GTTTACTAAT	ATTTTGTGTA	CAGAGTATAA	AGACCCCTATA	GTGGGTAAAT	TAGATACTAT	3120
	TAGCATATAA	TTAATTTAAT	GTCTTTATCA	TGGGATCTTT	TGCACTGCTTT	AATCTGGTTA	3180
15	ACATATTATA	ATTTGCTTTT	TTTCTCTTTA	CCTGAAGCCT	CTGTGTATAG	TATTTTCATGA	3240
	CATCGTTGTA	CAGTTTAACT	ATATCAATAA	AAAGTTTGGG	CAGTATTATA	ATATTGCAAA	3300
	TATGTTTAAAT	TATACAAATC	AGAATAGTAT	GGGTAATTAA	ATGAATACAA	AAAGAAGAGC	3360
	CTCTTTCTGC	AGCCGACTTA	GACATGCTCT	TCCCTTTCTA	TAAAGCTAGAT	TTTGAATATA	3420
	AGGGTTTCAG	TTAATTAATC	TATTTTCAGG	TTATGTCATC	TAACTTATAG	CAAACTACCA	3480
20	CAATACAGTG	AGTTCTGCCA	GTGTCACAGT	ACAAGGCATA	TTTCAGGTGT	GGCTGTGGAA	3540
	TGTAATAAATG	CTCAACTTGT	ATCAGGTAAT	GTTAGCAATA	AATTAATATG	TAGAATGAT	3600
	TAATCGGTGA	CATGTTACTG	TAATTAACCT	ATTGCACTTC	AAAACCTAAC	TTCCATCCTG	3660
	AAITTTATCAA	GATGTTCACT	ATTGTCATTT	GTTTGTGTTT	TATTGAAAG	TAAATGTTGTC	3720
	TTAAGATTATA	GAAGTGATTA	TTAGCTTGAG	AACTATTACC	CAGCTCTAAG	CAATAATGTA	3780
25	TTGTATACAT	ATTAAGATAA	TGGTTAAATG	CGGTTTACC	AAGTTTCCC	TTGAAATATG	3840
	AAITTCCTGTA	TGGGATTTA	TTGTGCAGCC	CTAAGCTTCC	TTCCCATTTT	ATGAATATAA	3900
	GGCTCTGATA	ATTGGACTGG	CAGGGGAAAG	AATGGTAGAG	ACAGAAATTA	AGACTTTATC	3960
	CTTGTGTTCT	TGTAATAACT	TATTTTCTTG	CTAATGTAAC	ATTTGTCTGT	TCCAGTGATG	4020
30	TAAAGATTAAT	AACTGATTTA	GCTAAATATT	AAITTTCAAA	AATAGTCCCT	CTTTAACTTA	4080
	GATATTTTCA	AGCTGGATTT	AGGAAGATCT	GTTATTTCTG	AAGTACTAAA	AAGAATAATA	4140
	CAAGGTACAA	TGCTGCTATT	CACATAATTCA	TGTTCCAGAA	GAGGAAATTA	TGAAGATATA	4200
	CTCAGTAGAG	TACTAGGTAG	GAGGATATGG	AAATTTGCTC	ATAAAATCTC	TTATAAAACG	4260
	TGCATATAAC	AAATGACAC	CCAGTAGGCC	TGCATTACAT	TTACATGACC	GTGTTTATTT	4320
35	GCCATCAAT	AACTGAGTAA	CTGACACCAG	ACAAAGACTC	CAAGTCATA	AAATAGCCTA	4380
	TGACCAACTG	CAGCAAGACA	GGAGGTCAGC	TCCCTATATA	TGGTGCTTAA	AGTGTGATTG	4440
	ATGTAATTTT	CTGTACTCAC	CATTTGAAGT	TAGTTAAGGA	GAACTTTATT	TTTTTAAAAA	4500
	AAGTAAATGG	CAACCACTAG	TGTGCTCATC	CTGAAGCTGT	ACTCCAATATC	CACTCCGTTT	4560
	TTAAAGCAAA	ATTATCTTGT	GATTTTAAGA	AAAGAGTTTT	CTATTTATTT	AAGAAAGTAA	4620
	CAATGCACTG	TGCAAGCTTT	CAGTAGTCTT	ATTGCTGCTAT	ATTGCTGCTAT	TAAACTCTT	4680
40	ACTACGTAAC	CAGTAATCAC	AAGGAAAGTG	TCCCTTTTGC	ATATTTCTTT	AAAATCTTTT	4740
	CTTTGGAAAG	TATGATGTTG	ATAATTAAC	TACCTTTATC	TGCCAAAACC	AGAGCAAAAT	4800
	GCTAAATACG	TTATGCTTAA	TCAGTGGTCT	CAAAATCGATT	TGCTCCCTTT	TGCTCCCTCT	4860
	GAGGGCTGTA	AGCCTGAAGA	TAGTGGCAAG	CACCAAGTCA	GTTTCCAAAA	TTGCCCTCTA	4920
45	GCTGCTTTAA	GTGACTCAGC	ACCCTGCTCT	AGCTTCAGCA	GGCGTAGGCT	CACCTGCGGC	4980
	GGAGCAAAAT	ATGGCCAGG	GAGAACTACA	GCTACGAAGA	CCTGCTGCTG	AGTTGAGAAA	5040
	AGGGGAGAA	TTATGGTCTG	AAITTTCTAA	CTGTCTCTTT	TCTTGGGTCT	AAAGCTCATA	5100
	ATACCAAAAG	GCTTCAGAC	CTGAGCCACA	CCGAGGCCCT	ATCTGGAACA	GGAGACTAAA	5160
	CAGAGGCMAA	TCAACCTTAG	GAAATACTTG	CATTCTGCCC	TACGTTAGT	ACCAGGACTG	5220
	AGGTCAATTC	TACTGAAAA	GATTTGTAGA	TTGAACTTAT	CTGATCGCTT	GAGACTCCTA	5280
50	ATAGGCAGGA	CTCAAGGCCA	CTAGAAAATT	GACAGTTAAG	AGCCAAAAGT	TTTTTAAATA	5340
	TGCTACTCTG	AAAAATCTCG	TGAAGGCTGT	AGGAAAAGGG	AGAACTCTCC	ATGTTGGTGT	5400
	TTTTCTCTGA	AAGATCAGTT	TGGGGTATGA	TATAAGCAGG	TATTAATAAA	AATAACACAC	5460
	CMAAGAGTTA	CGTAAACAT	GTTTTATTAA	TTTTGTGCCC	CACGTACAGA	CATTTTATTT	5520
	CTATTTTGAA	ATGAGTTATC	TATTTTCTAA	AAAGTAAAC	ACTATTAAG	TGCTGTTTTA	5580
55	TGTGAATAA	CTTGAATAT	GTCTCTATAA	AAATAGATC	ATAACTCATG	ATATGTTTGT	5640
	AATCATGTTA	ATTTAGATTT	TTATGAGGAA	TGAGTATCTG	GAAATATGTT	AGCAATACTT	5700
	GGTTTAAAT	TTTGGACCTG	AGACACTGTG	GCTGTCTAAT	GTAATCCTTT	AAAATTTCTC	5760
	TGCATTGTCA	GTAATGTAG	TATATTATTG	TACAGCTACT	CATAATTTTT	TAAAGTTTAT	5820
60	GAAGTTATAT	TTATCAATA	AAAACITTTCC	TATAT			

Seq ID NO: 315 Protein sequence:
Protein Accession #: XP_087254

65	1	11	21	31	41	51	
	MQFRECSING	MKYQEINGRL	VPEGPTPDSS	EGNLSYLSL	SHLNNLSHLT	TSSSFRTSPE	60
	NETELIKEHD	LFPAKAVSLCH	TVQISNVQTD	CTGDGPWQSN	LAPSQLEYVA	SSPDEKALVE	120
	AAARIGIVFI	GNSEETMEVK	TLGKLERYKL	LHILEFDSDR	RRMSVIVQAP	SSEKLLFAKG	180
70	AESSILPKCI	GGEIEKTRIH	VDEFALKGLR	TLCIAYRKPT	SKEYEIEDKR	IFEARTALQQ	240
	REEKLAAPQ	PIEKDLILLG	ATAVEDRLQD	KVREIEALR	MAGIKVWVLT	GDKHETAVSV	300
	SLSCGHFHT	MNILELINQK	SDSECAEQLR	QLARRITEDH	VIQHGLVVDG	TSLSLALREH	360
	EKLPMVCRN	CSAIVLCCRM	PLQKAKVIRL	IKISPEKPIT	LAVGDGANDV	SMIQEAHVGI	420
	GIMGKBGRQA	ARNSDYALAR	FKFLSKLLPV	HGHFYIIRIA	TLVQYFFYKN	VCFITPQFLY	480
75	QFYCLFSQQT	LYDSVYLTLY	NICFTSLPIL	IYSLLEQHVD	PEVLQNKPTL	YRDISKNRL	540
	SIKTFLYWTI	LGFSHAFFIF	FGSYLLIGKD	TSLGNGQMF	GNWTFGLTVF	TVMVITVTVK	600
	MALETHFWTW	INHLVWTGSI	IFYFVPSLFY	GGILWPPLGS	ONMYPVFIOL	LSSGSANFAI	660
	ILMVVTCFLF	DIHKVPDRH	LHPTSTZKAQ	LTETNAGIKC	LDSMCCFPFG	EACASVGRM	720
80	LERVIGRCSP	THISRSWSAS	DPFYTNDRSI	LTLSTMSST	C		

Seq ID NO: 316 DNA sequence
Nucleic Acid Accession #: NM_004473
Coding sequence: 661..1791

85	1	11	21	31	41	51	
	CTGCCAGCG	GTCCGCGGG	CTGGAGACCC	ACGCCGTGGA	GAGGACCAGC	CTCAGGTCGC	60

CCGGCTGGG CCGGCGCCCC GACCTCGCTG CCCCCGCTCT GCCTCTCTGC CCGTGGCGCT 120
 TACCGCCACC TTGGCTCTGG GGGCAGGGCA TGGGCGGCCC CCGCCAGATC GCCCAGCGCC 180
 AGTACTAACT GCCTCGCTC TGGCCTTCGA GCCCGAAGCC TCTTCTGCGC GCACAACCTA 240
 GGCAGTAATC CTAACACTAG GGGCACCACA GACCAGCTGC AGCCACCCCA ACCCAGGGAT 300
 5 CACTTCCGGA CCCCTCGACC GCCCGGCACC AGCGGCGAAG GGACCCCTCA GCGCGAGACC 360
 AGAGTCCAGT CCGGCTGGCG AGGCCACCGC CGCTGCCCCG CTCGAGAAGC ACAACGCGGG 420
 CTGAGCGCTG GGCCTAGCGG TCACTCCCCA GCCTCTGTCT GCACGCGGCC AGCCCCAGAC 480
 CACGAGCCTG GAGCTCCAG CGCGCGCCAG CCTGGGCGCG TGGGCTCTCC GGGCCAGCCC 540
 10 GCGACGATCC CCTGAGCTCT CCGCAGAAG GCGAGCGTC CGTTCGCGGG ACGCCAGGCC 600
 CGCCCCCGCC CCGGCACAGC CGCGGGGATC CAGAGCCCGG GGGTGCSSGA CGCCCGCGCC 660
 ATGACTGCGG AGAGCGGGCC GCCCGCGCGC CAGCGGAGG TGCTGGCTAC CGTGAAGGAA 720
 GAGCGCGCGC AGACGCGAGC AGGGGCGCGG GTCCAGGGG AGGCCACGGG CCGCGGGGCG 780
 GCGGCGCGCG CCGCAAGCG CCCCTGCGAG CGCGGGAAGC CGCCCTACAG CTACATCGCG 840
 15 CTCATGSCCA TGGGCATCGC GCACGCGCCC GAGCGCGGCC TCACTGTGGG CGGCATCTAC 900
 AAGTTTATCA CCGAGCGCTT CCCCTTCTAC CGCGACAACC CAAAAAAGTG GCAGAACAGC 960
 ATCCGCCACA ACCTTCACTT CAACGACTGC TTCTCTAAGA TCCCGCGCGA GGGCCGCGCC 1020
 CCGGTAAGG GCAACTACTG GCGCTCGAC CCCAACGCGG AGGACATGTT CGAGAGCGCG 1080
 AGCTTCTGTC GCGCGCGCAA GCGCTTCAAG CGCTCGGACC TCTCCACCTA CCCGCTTAC 1140
 20 ATGCAAGAGC CGCGCGCTGC CGCAGCCGCC CTGCGCGCAG CGCGCGCGCC CGCGCGCGCC 1200
 GCGGCGATCT TCCAGGCGCG GGTGCGCGCC GCGCGCGCCC CCTACCGCGG CGCGCTCTAT 1260
 GCAGGCTAGC CGCGCGCGCT GCTGCGCGCG CGCGCTCCAG TCTACTACCC CGCGCGCTCG 1320
 CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT 1380
 CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT CCGCGCGCTT 1440
 25 ACCACCGGCT ACCAGCGCGC AGGCTGCGCC GGGGCGCGCC CGGCCAACCC CTCTGCTTAT 1500
 GCGGCTGCTT ACAGCGCGCC CGACGCGCGC TACCGCGAGG GCGCGCGCAG TGCGATCTTT 1560
 GCGGCTGCTT GCGCGCGCTT GCGCGCGCTT GCGCGCGCTT GCGCGCGCTT GCGCGCGCTT 1620
 GTGGAGACCA CGGTGACTTT CTACGGGCGC ACCTGCGCGC GCGCGCTTGG AGCGCTGGGA 1680
 GCCTGCTACA ACCCTGCGCG GCAGCTCGGA GGGGCGCAGT CAGGCGCTTA CCATGCTCGC 1740
 30 CATGCTGCGC CTATCCCGG TGGGATAGAT CGGTTCGTGT CCGCATGTG AGCCAGCGTA 1800
 GGGAGCAAAA CTCATAGACA CATCGGCTGT TCACACGTTT CCGCAACCT GAGAACGAAC 1860
 AGGAATGGAG AGAGGACTCA ACTGGGACCC ACCTGGAAAA GACCGAGCAG GCCACAGAGG 1920
 CTGCTCTCC CCGCGCACAG CTAGGACACC CTGTGTACTT TGTAAACGGG AGGAGGTGGG 1980
 GCGAGGCGAC CAGAGCCCTT GGACTGCGAC AGGGACCTTC GATGGAGCGA AGCCCTCAA 2040
 35 CGGGATGCTT TCTGCGATTC TATCGGGGAG GGTCTTGGC GGTAAACGAG GGGCAGCGTA 2100
 GTGTCAACAC CAGAGACCAAG GATCCAAAT GTGGGGAATC AGTTTCAGCC TTCCATGTGC 2160
 TGCGGAACT CCGGCGCTTT TACCGGCTTC GTCTCTAGT GCCTTTAACT GCGTTACTAC 2220
 AATAAAGGC TGGCGCAGCG CCTTTCTTCT TAAAGTGAGG AGGACAAAT TGCAAAAGAA 2280
 ATAGGCTTTT CTCTTTTTT AAATGGAGA AATCTCTGCT CTGGTTGACC TGGGCTGGTT 2340
 40 TTCCCTGTCT CTGAGAATTT GAGACCTAGC TCCGAGTTGA ACTGTGCTC AGCACTCCAG 2400
 TCCATCATCA TCCATCATCA GTCTCCCGCA TCTGTATAC TAGAGGGCTG CAGGACTCTA 2460
 TCCACCGCCC CCGGCTTATC ATTGAGGGCC CCATCATCTT GGATGCTGCC CTGCGTATTT 2520
 GGCAGCAATG GTGGGCGACC CAGGGCCTCT GAGTAGCCAC CCAAGCCTA GCGCTGTTT 2580
 TAGGGAACGG AAAAGAGTTC ATGGCCAAGC GTCTAACCTA AAGTCCAGG ATTGGCTCCA 2640
 45 GGCAGCAATT ATATCATAC TTATTGAAT TTTGAGCAGG ACGTGTGGT AATTTCATGG 2700
 CTGTACTGCG CCGATCATAA ATCTGCTTTT CCATTATAAG CGAGAGAGAA GTACATTCGT 2760
 TCATTTGTCC ACTGTTCTTT GTCTACGCG AGCCCTGGAC CCAAGGGTG AACTAAAGTT 2820
 TAAGSAGATG AGAGSATTCA AGGAGCCCGT TGGTGAAGCC TTTCACTAGC TGGGAGGGC 2880
 TCTTCCATCC CCGACACCCC CTGCTACACC TCAGCAGCCT CCCCCATGCA AAAAGGAAG 2940
 50 AGAAAAATTA AGTTAGGGCA GTGAGTAAAG TGAGCTTTAG AAAGAAACTG GAATTTTAAC 3000
 TTCAATTTGT ATCTTGCTTA AGTAGCAGGC TCACTAAAAT TAGAGAAAGT CCAATAACTC 3060
 TCCCCCTTTC CCTTAGAGAA TCTTTAAGTT TCGATTCTGG AGCAAAACTT TTCAGCATTA 3120
 AATATTTCAG AGGCTCCATT CACAGCTTTC AGATAAACTG GAGTGTTCAG ATGGACTGTT 3180
 TTAATAAAAA TCTTTGAGCA AGTGAGTTAT GGCAGAGAGAA ACTCAGCCTC TTTCTGTATA 3240
 55 AACTTAACAG GGAAGGGCTG GGGTGTGAAA AAGAAGATTG TATGAAAACC ATTGGTAATT 3300
 TTTATTTTGT ATTTTGGGA CTGCACTATC CTGTTTACGA AGACATGTGA ACTTGGTTCA 3360
 GTCCAAATGG GGAATTTGTAT AAACCACTGC TCTCCATTAG AAATATGGTG CAAGCCACAT 3420
 ATGTAATTTT AAATATTCTA GTAGCCACAT TAATAAAGTN AAAAGAAACA AAAAAAATA 3480
 AA

60 Seq ID NO: 317 Protein sequence:
 Protein Accession #: NP_004464

1 11 21 31 41 51
 65 FKHLTHYROI DTRANSCRIP TIONFACTOR TIFMTAESGP PPPQPEVLAT VKEERGETAA 60
 GAGVPGEATG RGAGRRRRKR FLQRKPPYS YIALIAMAIA HAPERRLTG GIYKFITERF 120
 PFYRDNPKKW QNSIRHNLTL NDCFLKIPRE AGRPGKGNWY ALDPNAEDMF ESGSPLRRK 180
 RFKRSLDSTY PAYMEDAAAA AAAAAAIAA AAAAAIIPGA VPAARPPYPG AVYAGYAPPS 240
 70 LAAPPPVYYP AASPGPCRVF GLVPERPLSP ELGPAPSGPG GSCAFASAGA PATTTGYQPA 300
 GCTGARFANP SAYAAAYAGP DGAYPQAGS AIFAAAGRLA GPASPPAGGS SGGVETTVDF 360
 YGRTPSQPGF ALGACYNPGG QLGGASAGAY HARHAAAYPG GIDRFVSAM

75 Seq ID NO: 318 DNA sequence
 Nucleic Acid Accession #: NM_005688
 Coding sequence: 126..4439

1 11 21 31 41 51
 80 CCGGCGAGGT GGCCTATGCT CGGAGCGGTG GTTGAGCGGC TGGCGCGGTT GTCCTGGAGC 60
 AGGGGCGCAG GAATTCGTAT GTGAACTAA CAGTCTGTGA GCCCTGGAAC CTCGCTCAG 120
 AGAAGATGAA GGATATCGAC ATAGGAAAG AGTATATCAT CCCCAGTCTT GGGTATAGAA 180
 GTGTGAGGGA GAGAACACAG ACTTCTGGGA CGCACAGAGA CGGTGAAGAT TCCAAAGTTCA 240
 85 GGAGAACTCG ACCGTGGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCGA GCCAGGGGCC 300
 TCTCTCTTGA TGCTTCCATG CATTCTCAGC TCGAATCTCT GGATGAGGAG CATCCCAAGG 360
 GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACCCAGC 420
 ACCCAGTGA CAATGCTGGG CTTTTCCTCT GTATGACTTT TTCGTGGCTT TCTTCTCTG 480

	CCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
	AAGTTGGGCC	AGACGCTGCT	TCCCTCGGAA	GGGTGTGTGT	GATCTTCTGC	CGCACCAGGC	660
5	TCACTCTGTC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCACT	GAACACAGCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTTAGT	GCTGGGCTCT	CTCTGACGCG	AAATCGTGGG	GTCTTGGTGG	CTTGCACTGA	840
	CTTGGGCAAT	GAATTACCGA	ACCGGTGTCC	GCTTGCGGGG	GGCCATCCTA	ACCATGGCAT	900
	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
10	TTTGCTCCAA	CGATGGGCGA	AGAATGTTTG	AGGCAGCAGC	CGTGGCAGC	CTGTGGCTGG	1020
	GAGGACCCGT	TGTTGCCATC	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCCTGGG	ATCAGCTGTT	TTTATCCTCT	TTTACCAGC	AATGATGTTT	GCATCACGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CGGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTCTCT	1260
15	AGATGTGTTA	AAATATCCGC	GAGGAGGAGC	GTGCGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGTATCAC	TGTGGGTGTG	GCTCCCATTG	TGGTGGTGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTTCATAT	GACCTCGGGC	TTOGATCTGA	CAGCAGCACA	GGCTTTCACA	GTGGTGACAG	1440
	CTTTCAATTG	GATCAGCTTT	GCTTTGAAAG	TAAACCGGTT	TTCACTAAAG	TCCCTCTCAG	1500
	AAGCCTCAGT	GGCTGTGTAG	AGATTTAAGA	GTTTGTCTCT	AATGGAAGAG	GTTCACATGA	1560
20	TAAAGAACAA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAATAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCACTCCAGT	ATCCAGAACT	CGCCCAAGCT	GACCCCAAAA	ATGAAAAAAG	1680
	ACAAGAGGGC	TTCCAGGGGC	AAGAAAGAGA	AGGTGAGGCA	GCTGCAGCGC	ACTGAGCATC	1740
	AGGCGGTGCT	CGCAGAGCAG	AAAGGCCACC	TCTCTCTGGA	CAGTGAACGAG	CGGCCAGTGC	1800
	CCGAAGAGGA	AGAAGGCCAG	CACATCCACC	TGGGCCACCT	GGCTTACAG	AGGACACTGC	1860
25	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AACCTGGTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCATT	TCAGCCATTT	TAGGCCAGAT	GACGCTCTTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACCTCT	GCTTATGTGG	CCCAGCAGGC	CTGGATCTCT	AATGCTACTC	2040
	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTG	CCTGAGGCTT	GACCTGGCCA	TTCTTCCCAG	CAGCGACCTG	ACGAGATTGT	2160
30	GAGAGCGAGG	AGCCAACTGT	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCGGGCCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTT	CAGTGCTTTA	GATGCCCATG	2280
	TGGCAACACA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CGAGTTACAG	TACCTGGTGT	ACTGTGATGA	AGTGATCTTC	ATGAAAGAGG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACGTATGAA	TTTAAATGGT	GACTATGCTA	2460
35	CCATTTTTAA	TAACTGTGTT	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCCAGTG	TTACAGAGAG	AAGTCACAAG	ACAAGGGTCC	TAAACAGGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCAAGTCC	CTGGTCAGTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
	TCCTGGTTAT	TATGGCCCTT	TTTATGCTGA	ATGTAGGCGG	CACCGCCTTC	AGCACCTGGT	2760
40	GGTTGAGTTA	CTGATCAAG	CAAGGAAGCG	GGAAACCCAC	TGTGACTCGA	GGGAACGAGA	2820
	CTCGGTGTAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCG	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCTCTA	AAGCCATTGG	AGGAGTGTGC	TTTGTCAAGG	2940
	GCACGCTGGC	AGCTTCTCTC	CGGCTGCATG	ACGAGCTTTT	CGGAAGGATC	CTTGAAGGCC	3000
	CTATGAAGTT	TTTTGACACG	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
45	TGGATGAAGT	TGACGTGGGG	CTGCCGTTC	AGGCCGAGAT	GTTCATCCAG	AACTTATGCC	3120
	TGGTGTCTCT	TTTACAGTGA	ATGATGCGAG	GAGTCTTCCC	GTGGTCTCTT	GTGGCAGTGG	3180
	GGCCCTCTGT	CATCCTCTTT	TCAGTCTGTC	ACATTGTCTC	CAGGGTCTCT	ATTCTGGGAGC	3240
	TGAAGGCTCT	GGACAATATC	ACGCAATCAC	CTTCTCTCTC	CCACATCACG	TCCAGCATAC	3300
	AGGGCCTTGC	CACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTCCTGCAC	AGATACCAGG	3360
50	AGCTGTCTGA	TGACAAACCA	GCTCTTTTTT	TTTTGTTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGGGGCT	GGACATGCATC	AGCATCGCCC	TCATCACACC	CACGGGGCTG	ATGATCGTTC	3480
	TTATGACCGG	GCAGATTCCC	CCAGCCTATG	CGGCTCTGCG	CATCTCTTAT	GCTGTCCAGT	3540
	TAACTGGGGT	GTCCAGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTACCTT	3600
	CGGTGGAGAG	GATCAATCAC	TACATTAAGA	CTCTGTCTCT	GGAAAGCACC	GCCAGAAATTA	3660
55	AGAACAGAGC	TCCCTCCCCC	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACCGCA	3720
	AGATGAGGTA	CCGAGAAAAA	CTCCCTCTTG	TCTTAAAGAA	AGTATCCTCT	ACGATCAAAAC	3780
	CTAAAGAGAA	GATTGGCAAT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTCTCG	TCTGGTGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAATCA	3900
	GTGATATTGG	CCTTGGCGAC	CTCCGAAGCA	AACCTCTCTAT	CATTCTCTCA	GAGCCGGTGC	3960
60	TGTTTCACTG	CATGTCGAGA	TCAAATTGAG	ACCCCTTCAA	CCAGTACACT	GAAGACGAGA	4020
	TTTGGGATCG	CTGGGAGAGG	ACACACATGA	AGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TTGAATCTGA	AGTATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
65	CCATGGACAC	AGAGACAGAC	TTATTGATTC	AAGAGACCAT	CGAGAGAAGCA	TTTGCAGACT	4260
	GTACCATGCT	GACCATTGCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTCTCTG	TCCAACGACA	4380
	GTTCGCCGAT	CTATGCCATG	TTTGTGCTG	CAGAGAACAA	GGTCTCTGTC	AAGGGCTGAC	4440
	TCTCTCCCTG	TGACGAAGTC	TCTTTTCTTT	AGAGCAITGG	CATTCCCTGC	CTGGGGCGGG	4500
70	CCCTCATGCG	CGTCTCTCTA	CGAAACCTTT	GCCCTTCTCG	ATTTTATCTT	TGCGACAGCA	4560
	GTTCGGGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTGTATTT	4620
	ATTCCATATT	CATGTAAACA	AAATTAGTGT	TTTGTCTCTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAAATT	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
75	TATTAAAAATA	AGCAGTGTGC	TAATAACAGT	GCATATTCTT	TTCTATCATT	TTTGTACAGT	4860
	TTGCTGTACT	AGAGATCTCG	TTTTGCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCACTCTT	4920
	CTCTAGCTGG	TGGTTTCAAG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCCTCTGCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
	GAGACGGGTG	GGCGGCTGGA	GACCATGCAG	AGCGCGGTGA	GTCTCAGGGG	CTCTGCGCTT	5100
80	CTGTCTGTGT	GTCACTTACT	GTTCCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCCC	5160
	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCTGC	TCTTCTTTTT	GCTGTGTTTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCATCTGCC	TCAGGTTCTT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340
	GTGGTTTCCA	AGCCCTGGAG	CCAAGTCTGT	CTTTTGTAGG	TGGCACTTTT	TCAATTGCTT	5400
85	ATTCCACAC	CTCCACAGTT	CAGTGGCAGG	GCTCAGGATT	TGCTGGGTCT	GTTCCTCTTT	5460
	CTCACGCGAG	TGCTCGCACA	GTCTCTCTCT	CTCTCTCCCC	TCAAAGTCTG	CAACTTTAAG	5520
	CAGCTCTTGC	TAACTCAGTT	CTCACACTGG	CGTAGAAGTT	TTTGTACTGT	AAAGAGACCT	5580
	ACCTCAGGTT	GCTGGTGTCT	GTGTGGTTTT	GTGTGTTCCC	GCAAAACCCC	TTTGTGCTGT	5640
	GGGGCTGGTA	GCTCAGGTGG	CGGTGGTCAC	TGCTGTCTATC	AGTTGAATGG	TCAGCGTTGC	5700

ATGTCGTGAC CAACTAGACA TTCTGTCGCC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760
 CAAAAATCTG AAAATGTGAA TAAAAATTATT TTGGATTGTT TAAAAAATAA AAAAAAATAA 5820
 AAAAAAATAA AAAAAAATAA

5

Seq ID NO: 319 Protein sequence:
 Protein Accession #: NP_005679

10	1	11	21	31	41	51	
	MKOIDIGKEY	IIPSPGYRSV	RERTSTSGTH	RDREDSKFRF	TRPLECQDAL	ETAARAEGLS	60
	LDASMHSQLR	ILDEEHPKKG	YHHGLSALKP	IRTTSKHQHP	VDNAGLPSCM	TFSWLSLAR	120
	VAHKKGELSM	EDVNSLSKHE	SSDVNCRRL	RLWQEEELNEV	GPDAASLRV	WMIFCRRLI	180
	LSVICLMITQ	LAFPSGPAFM	VKHLEYTQA	TESNLQYSL	LVLGLLLEI	VRWSLALTW	240
15	ALNYRTGVRL	RGAILTMAFK	KILKLNKKE	KSLGELINIC	SNDQGRMFEA	AAVGSLLAGG	300
	PVVAILGMIV	NVILIGPTGF	LGSAPVILFY	PAMMFASRLT	AYFRKCVAA	TDERVQKME	360
	VLTYIKFIKM	YAWVKAFSQS	VQKIREEER	ILEKAGYFQG	ITVGVAIVV	VIASVVTFSV	420
	HMTLGFDLTA	AQAFVTVTVF	NSMTFALKVT	PFSVKSLSEA	SVAVDRFKSL	FLMEEVHMIK	480
	NKPASPHIKI	EMKNATLAWD	SSHSSIQNSP	KLTPKMKKDK	RASRGKKEKV	RQLQTEHQA	540
20	VLAEQKGLHL	LDSDERPSP	EEBEGKIHIL	HLRLQRTLHS	IDLEIQEGKL	VGICGSVGS	600
	KTSLISAILG	QMTLLGESSA	ISGTFAYVAQ	QAWILNATLR	DNILFGKEYD	EERYNSVLNS	660
	CCLRPDALIL	PSSDLFEIGE	RGANLSGGQR	QRISLARALY	SDRSIYILDD	PLSALDAHV	720
	NHIFNSAIRK	HLKSKTVLTV	THQLQYLVD	DEVIFMKEGC	ITERGTHEEL	MNLNGDYATI	780
	FNLLLGGETP	PVEINSKKET	SGSQKKSQDK	GPKTGSVKKE	KAVKPEEGQL	VQLEEKQGS	840
25	VPSVYGVYI	QAAGPLAPL	VIMALFMAN	GSTAPSTWML	SYWIKQSGSN	TTVTRGNETS	900
	VSDSMKDNFH	MQYIYASIAL	SMAMVLLIKA	IRGVVVFVKT	LRASSRLHDE	LFRRILRSPM	960
	KFEDTTPFGR	ILNRPFSKMD	EVDVRLPFOA	EMFIQNVILV	FFCVGLIAGV	FWFLVAVGP	1020
	LVILFSLVHI	VSRVLIRELK	RLDNITQSPF	LSHITSSIQG	LATIHAYNKG	QEFLEHRYQEL	1080
	LDNQAPFFL	FTCAMRWLAV	RDLISIALI	TTTGLMIVLM	HQGIIPYAG	LAISYAVQLT	1140
30	GLPQFTVRLA	SETEARFTSV	ERINHYIKTL	SLEAPARIKN	KAPSPDWPE	GEVTFENAE	1200
	RYRENPLVPL	KYVFTPIKPK	EKIGIVGRTG	SGKSSSLGMA	FRVLVSGGC	IKIDGVRISD	1260
	IGLADLRSLK	SIIPQEPVLF	SGTVRSNLD	FNQYTEDQIN	DALETHMKE	CIADLPKLE	1320
	SEVMENGDNF	SVGERQLLCI	ARALLRHCKI	LILDEATAAM	DTETDLTIQE	TIREAFADCT	1380
35	MLTIAHRLHT	VLGSDRIMVL	AQGQVVEFDT	PSVLLSNDSS	RFYAMFAAAE	NKVAVKG	

Seq ID NO: 320 DNA sequence
 Nucleic Acid Accession #: AK022089.1
 Coding sequence: 181-1488

40

40	1	11	21	31	41	51	
	AGCAGTTGCA	CAACTTCCAG	CAACTTTCCTC	AGCCGGCTAC	TAATGAGCTG	AAAGCCAGGA	60
	ACATCCGAGG	AGAAGAGAAA	GCTTCCAGCC	CTCTCCCTT	CACCTCGGAA	ATCCAGACAC	120
45	CCCCACCCCC	AACTCCAGAT	CACCTTAAGA	TAATTTCTTT	ATTCGTTTGC	CCGACAGACC	180
	ATGGCTCCCT	TTGGAAGAAA	CTTGCTAAAG	ACTCGGCATA	AAAACAGATC	TCCAACTAAA	240
	GACATGGAAT	CAGAAGAGAA	GGAAATTTGT	GTTTGGGTTT	GCCAGAAGA	GAAGCTGTGC	300
	TGTGGGCTGA	CTAAAGCCAC	CACCTCTGCT	GATGTCATCC	AGGCTTTGCT	TGAGGAACAT	360
	GAGGCTACGT	TTGGAGAGAA	ACGATTTCTT	CTGGGGAAGC	CCAGTGATTA	CTGCATCATA	420
50	GAGAAGTCCA	GAGGCTCCGA	AAGGGTTCTT	CCTCCACTAA	CTAGAATCCT	GAAGCTTTGG	480
	AAAGCGTGGG	GAGATGAGCA	GCCCAATATG	CAATTTGTTT	TGTTAAAGC	AGATGCTTTT	540
	CTTCCAGTTC	CTTTGTGGCG	GACAGCTGAA	GCCAAATTAG	TGCAAAACAC	AGAAAATTG	600
	TGGGAGCTCA	GCCAGCAGAA	CTACATGAAG	ACTTTACCAC	CAGATAAACA	AAAAAGARTA	660
	GTCAAGAAAA	CTTTCCGGAA	ACTGGCTAAA	ATTAAGCAGG	ACACAGTTTC	TCATGATCGA	720
55	GATAATATGG	AGACATTAGT	TCATCTGATC	ATTTCCAGG	ACCATACTAT	TCATCAGCAA	780
	GTCAAGAGAA	TGAAGAGCT	GGATCTGGAA	ATTGAAAGT	GTGAAGCTAA	GTTCCATCTT	840
	GATCGAGTAG	AAAAATGATG	AGAAAATAT	GTTCAGGATG	CATATTTAAT	GCCAGTTTC	900
	AGTGAAGTTG	AGCAAAATCT	AGACTTGCG	TATGAGGAAA	ACCAGACTCT	GGAGGACCTG	960
	AGCGAAAGTG	ATGGAATTGA	ACAGCTGGAA	GAACGACTGA	AATATTACCG	AATACTCATT	1020
60	GATAAGCTCT	CTGCTGAAAT	AGAAAAAGAG	GTAAGAAAGT	TTTGCAATGA	TATAAATGAA	1080
	GATCGCGAAG	GGGAGCTGCG	AAGTGAAGT	GAAAGCTCTA	ATTAGAGAG	TGTTAAGTGT	1140
	GATTTGGAGA	AAAGCATGAA	AGCTGGTTTG	AAAAATCACT	CTCATTTGAG	TGGCATCCAG	1200
	AAAGAGATTA	AATACAGTGA	CTCATTGCTT	CAGATGAAAG	CAAAAGAATA	TGAATCCTG	1260
	GCCAAGGAAT	TCAATTCAC	TCACATTAGC	AACAAAGATG	GGTGCCAGTT	AAAGGAAAC	1320
65	AGAGCGAAGG	AATCTGAGGT	TCCAGTAGC	AATGGGGAGA	TTCTCCCTT	TACTCAAAGA	1380
	GTATTTAGCA	ATTACACAAA	TGACACAGAC	TCGGACACTG	GTATCAGTTC	TAACCAAGT	1440
	CAGGACTCCG	AAACAACAGT	AGGAGATGTG	GTGCTGTTGT	CAACATAGTT	CCAATGGCTC	1500
	CTTCTGACC	TGCTTTTCATG	TTTAAATGTT	TGTTAATTT	AATAGGAAAC	CTCATTTTAA	1560
70	ATATAACACT	CAAAAAAATG	TAAATCATAT	TGTAGTATTC	AATAGTTAAT	AAAAACTCGA	1620
	GAAATGTGTT	GTTTCTG					

Seq ID NO: 321 Protein sequence:
 Protein Accession #: NP_005438.1

75

75	1	11	21	31	41	51	
	MAPFGRNLLK	TRHKNRSPK	DMDSEEKEIV	VWVQEEELV	CGLTKRTTSA	DVIQALLEEH	60
	EATFGKERFL	LGPSPDYCII	EKWRSERVL	PPLTRILKLN	KAWGDEQPM	QFVLVKADAF	120
80	LPVPLWRTAE	AKLVQNTKEL	WELSPANMYK	TLPPDKQKRI	VRKTRFKLAK	IKQDVTSHDR	180
	DNMETLVHLI	ISQDHTHQ	VKRMKELDLE	IEKCEAKFHL	DRVENDGENY	VQDAYLMPFS	240
	SEVEQNLDLQ	YEENQTLDEL	SESDGIBQLE	ERLKYRILII	DKLSAEIEKE	VKSVCIDINE	300
	DABGEAASEL	ESSNLESVKC	DLEKSMKAGL	KIHSLSGIIQ	KEIKYSDSL	QMKAEYELL	360
	AKEFNSLHIS	NKDGCCQLEN	RAKESEVPSS	NGEIPPTQR	VPSNYNTD	SDTGISSNHS	420
85	QDSETVGDDV	VLLST					

Seq ID NO: 322 DNA sequence
 Nucleic Acid Accession #: NM_030920.1

Coding sequence: 317-1123

	1	11	21	31	41	51	
5	AGCATTGAAG	GGGAAGGAAC	TGCGGGGTGTG	GTGTGTGTGTAT	GTGTGTGTGTGT	ATGTGTGTGTGC	60
	GGCGCGTGGG	TGCGTGTGTG	TGCGCGGCGCT	AGTGTGTGTGA	CAAGGAGGTG	GGGGCAGCTG	120
	AGTTAGAGTC	CCAACTCTTG	GACTCCATT	GCTATTCTCT	TCTTCTCC	CCACACCTAT	180
	CTGGTGGTGG	TAGTGGGGGT	TTATATTTC	GTTCTTTTC	ATTCAATTCT	AAATCTCTTA	240
	AAATATTTCG	GTGGGGGTG	TGGGGGAAGG	CAGGAAGGG	AAAAGGAGAG	TAGTAGCTGA	300
10	AGAGCAAGAG	GAGGACATGG	AGATGAAGAA	GAAGATTAA	CTGGAGTTAA	GGAAACAGATC	360
	CCCGGAGGAG	GTGACAGAGT	TAGTCTTGA	TAATTGCTG	TGTGTCAATG	GGGAAATTGA	420
	AGGCCTGAAT	GATACTTTCA	AAGAACTAGA	ATTTCTGAGT	ATGGCTAATG	TGGAACCTAAG	480
	TTGCTGGGCC	CGGCTTCCCA	GCTTAAATAA	ACTTCGAAAA	TTGGAGCTTA	GTGATAATAT	540
	AAATTTCTGA	GGCTTGAAG	TCCTGGCAGA	GAAATGTCCA	AATCTTACCT	ACCTCAATCT	600
15	GAGTGAAGAA	AAATATAAAG	ATCTCAGTAC	AGTAGAAGCT	CTGCAAAATC	TAAAAAATTT	660
	GAAAACTCTT	GACCTGTTTA	ACTGTGAGAT	CACAAACCTG	GAAGATTATA	GAGAAAGTAT	720
	TTTTGAACCT	CTGCAACAAA	TCACATACTT	AGATGGATT	GATCAGGAGG	ATAATGAAGC	780
	GCGGACTCT	GAAGAGGAGG	ATGATGAGGA	TGGAGATGAA	GATGATGAAG	AGGAAGAGGA	840
	AAATGAAGCT	GGTCCACCGG	AAGGATATGA	GGAAGAGGAG	GAGGAAGAGG	AAGAGGAGGA	900
20	TGAGGATGAG	GATGAAGATG	AAGATGAAGC	AGGTTGAGAG	TTGGGAGAGG	GAGAAGAGGA	960
	AGTGGGCTCT	TCATACTTAA	TGAAAGAAGA	AATTGAGGAT	GAAGAAGATG	ATGATGACTA	1020
	TGTTGAAGAT	GCGGAAGAGG	AGGAAGAAGA	GGAAGAAGGA	GGTCTTCGAG	GGGAGAAGAG	1080
	GAAACGAGAT	GCTGAAGAGC	ATGGAGAGGA	AGAAGATGAC	TAGATCATTC	TAAGACCAGA	1140
	TTCTCTAATG	TTTCTGGGTG	TGCAATAGAG	TGATCACATC	TTTGTTCCTT	CATGTACGAT	1200
25	AGCTATCCCT	ACGTAAGATA	ATGTGTAACT	TTTTATAGGA	AAAGTGTGGT	TTTACTATTT	1260
	TTGCTTATCT	ATTTCAAAAT	AGAACTAGTC	TGTTAATGAT	CATATTGTAT	GTAGAGAAAA	1320
	ATTTTCAATG	ACTCCCATTT	TGGAATTCCT	TAGCAATTTA	TTTAGACTTA	ATTTTATAAA	1380
	TTCAAGCTTA	CTGTATTAGT	CATTTTATAG	CCATAATTAA	AACATGATCA	CTTTTAAACA	1440
	GGTGTAGTAT	GGTGCATTTT	ATTCCTTAT	TATAGATTAA	CTGAAATTAC	AGTTTGTCTAT	1500
30	AAATATAAAT	GACAAATAGT	TCCTGAGTGG	TAAGTTGGTT	ATTTTTFAG	AGGTGATCCA	1560
	GGAATCTTTA	GTTTGAAGGC	AGTTACCTTT	TTTTTTTTTT	TTTTTTTTTG	ACTAAGAGTG	1620
	TTTGGTTGCT	TTTTGTGAC	AAGTAACTTG	GAAATAGAAA	GCAGAAATAG	AAAGGTTCTA	1680
	TTCAAGCAAC	TAGTTCATGG	ATTTTGTGGA	GGTCTTATTC	AGTAAATATG	TTCATGGATT	1740
	TAGTGTGTGAC	TGATAAGATT	TTATTTTGA	AGGAAAAATT	GCTTATACTA	AGTCCAGAGA	1800
35	CATGCAGGTG	AGCCCTTTTG	TCAGGCTGCA	AATCATGACA	TGCCGATGGT	TGTTTATTTT	1860
	GTCTTTAGGT	GTGCATCTCT	TTTCTTCTTA	GCAATTCCTT	TATGATCACC	TTCCCTTCTT	1920
	GTCTCACTCC	CTCCCGCTCT	CTCAAAAGGA	ACTTGGGAAA	CTTGTGAAAC	CCAGGAAAAAC	1980
	CTTTAGTCTT	ATACCTCAAC	TACGTTTCAG	TCCGTCTCTG	GTTTTAAATA	AGTGAAGTAG	2040
	AAGAAATTGA	GTAATTTCTG	ACATAAGAA	ATATTATCAA	TACAGTTTAA	TGCAGTAAGC	2100
40	TCTCCTTACC	ATAAATGTTT	CTTGGTTGAC	AACATCTAAG	ACAATATTAG	TGGGATGAAG	2160
	AAGAAAAAGC	AGGGGTGCTT	TTGGAAGCAG	TGTTAGTGTT	OCTCAAAAGT	CGGAACAAAT	2220
	GCCTGTTGAT	ATATTATAA	GACATTAAAG	TCAAAATTTA	ATGTTGGCCT	CTCAATGAT	2280
	TTGGATACCA	CTCTGCAAAG	TATTTCTAAC	CTTTAATTCC	CAGTTTAAAA	ACAGATATAA	2340
	TAATAGCAT	TAATTTGAAT	ATACTAGGCA	GCTGGAAAAA	TATTTGAAAC	TAAATTGACA	2400
45	TTAAATTTAA	TTTGTGTTT	CAAGTGGATG	TCCATTAAAA	GTAGAAAAAT	ATTTGGGATA	2460
	AGTGAGTGTG	TGTTTCTCTA	CATGGCTACT	AAATAAAAATA	TAATGAGTAT	ACAAGTATAT	2520
	CTCCTCTTTT	GCTATGGAGG	CTCCATGTTT	AAGGCAATGG	CTTTTAAAT	CTTGGCTATC	2580
	TAAAAATTTT	TCCCTTGTGT	TGGAATATTT	GTAAGTTT	AAGAAGTTAG	TGTGAGCAAA	2640
50	TTAATTTGAG	TTATGCTTCT	ATACTGGGAC	ATATTTAAAT	ACTGAGTATA	GTACTGCTGC	2700
	TACTGCTTCT	ACAAATGAAA	ATGTATGACT	TGGTGTTTTA	AAGTAAAAAT	TATGATGTTA	2760
	CTTGTGGAGA	AACATAAAAT	GTGTACAAC	TGACCGAAAG	AAAACCCCTG	GGGATAAGTT	2820
	TAGTGAGGGG	ATTGGAATCC	CCAAAAAGAT	AACATTTTTC	TCTGCTTTT	AAAAACTGAA	2880
	ATTCCCTGTT	CTAGTCTCTA	ACAAATCTCA	TTACATACTA	TGCCAGATTA	CAAAATACCT	2940
	ATTTTAAAAA	TGAAATCTAT	ATATTGACTT	TCTTATCAAT	CATCTTACTG	TGCAATCAAA	3000
55	ATTAGATGAC	TTTGGTTTGA	AAACAACACT	TAGAGCCTCC	AGATAACTTT	TAAGACTTAT	3060
	TTAGCTTTGT	GGGTGTGTAT	TTCAATGCAA	TAAGTAAGGG	TGGGTTTAT	ATTTGTAGTA	3120
	AGTTTTCGGT	CCTATTTTAA	TGCTCTTTGT	ATGGCAGTAT	GTATATATTG	TGTTAAGTTC	3180
	CTCAAGAATC	CTTTAAGATT	AATACTTTTG	TGCAACTGTG	TTTTGAATAA		3240
60	AGCCATGACA	GTGTTAAAAA	CAAAAC				

Seq ID NO: 323 Protein sequence:
Protein Accession #: NP_112182.1

	1	11	21	31	41	51	
65	MEMKKKINLE	LRNRSPEEVT	ELVLDNCLCV	NGEIEGLNDT	FKELEPLSMA	NVELSSLARL	60
	PSLNKLRLKE	LSDNIIISGL	EVLAEXCPNL	TYLNLGNKI	KDLSTVEALQ	NLKNLKSIDL	120
	FNCEITNLED	YRESIFELLQ	QITYLDGFDQ	EDNEAPDSEE	EDDEEDGEDD	EEEEEENBAGP	180
70	PEGYEEEEEE	EEEEDEDEDE	DEDEAGSELG	EGEEEEVGLSY	LMKEEIQDEE	DDDDYVEEGE	240
	EEEEEEBGL	RGEKRRKDAE	DDGEEEDD				

Seq ID NO: 324 DNA sequence
Nucleic Acid Accession #: NM_003812
Coding sequence: 224..2722

	1	11	21	31	41	51	
75	TCCTCTGCGT	CCCGCCCCGG	GAGTGGCTGC	GAGGCTAGGC	GAGCCGGGAA	AGGGGGCGCC	60
	GCCAGCCCCC	GAGCCCCCGG	CCCCGTGCCC	CGAGCCCGGA	GCCCCCTGCC	CGCGCGGCGA	120
80	CCATGCGCGC	CGAGCGCGCG	TGACCGGCTC	CGCCCGCGGC	CGCCCCCGAG	CTAGCCCGGC	180
	GCTCTCGCGG	GCCACACGGA	GCGCGCGCCG	GGAGCTATGA	GCCATGAAGC	CGCCCGGCGA	240
	CAGCTCGCGG	CAGCGCGCCC	TGGCGGGCTG	CAGCCTTGCC	GGCGCTTCTC	GCGGCCCCCA	300
	ACGCGCGCCC	GCGCGCTCGG	TGCTTGCCAG	CGCCCGCGGC	CGCACGCGCG	CCTGCGCCCT	360
	GCTTCTCGTC	CTTCTCTGCG	TGCTTCCGCT	CGCCGCTCGG	TCCCGGCCCC	GCGCTGGGGG	420
85	GGCTGCTGCG	CCGAGCGCTC	GCATTGGAA	TGAAACTGCA	GAAAAAATTT	TGGGAGTCTC	480
	GGCAGATGAA	GACAATACAT	TGCAACAGAA	TAGCAGCAGT	AATATCAGTT	ACAGCAATGC	540
	AATGCAGAAA	GAAATCACAC	TGCCTTCAAG	ACTCATATAT	TACATCAACC	AAGACTCGGA	600

Seq ID NO: 325 Protein sequence:
Protein Accession #: NP_003803

Seq ID NO: 326 DNA sequence
Nucleic Acid Accession #: AK074418.1
Coding sequence: 244-1515

308

	CGGAAAGCC	AGCTACATAA	GAAACGGGAA	GATGCGGAGT	TTTGGATGTC	GTGTCAAGAT	1200
	TTCCAACAGA	AATTCATCGC	CATGTTTATA	TGTAGCGAAA	TTCCAATTAC	CCTGGACCAT	1260
	GGAAACACAC	TCCACGAAGG	ATGGTCCCAA	ATAATGTTTA	GGAAAGCAAGT	GATTCTAGGA	1320
5	AACACTGCGAG	GAGGACCTCG	GAATGATGCT	CAATTCAACT	TCTCTGTGCA	AGAGCCCAATG	1380
	GAAGGCACCA	ATGTTGTGCT	GTGCGTCACA	GTGTCTGTCA	CACCATCAAA	TTTGAAGCA	1440
	GAAGATGCAA	AATTTCCACT	CGATTTCCAA	GTGATTCTGG	CTGGCTCACA	GAACACTGT	1500
	CCAAAGCTCA	AATAATAAAT	TCGCGCGCAA	CTTCACCATG	ACTTACCATC	TGAGCCCTGG	1560
	GAACATGTT	GTGGTTGCAC	AGACACGGAG	AAAATCAGCG	GAGTTCCTGC	TCCGAATCTT	1620
10	CCTGAAAATG	CCAGACAGTG	ACAGGCACCT	GAGCAGCCAT	TTCAACCTCA	GAATGAAGGG	1680
	AAGCCCTTCA	GAACATGGCT	CCCAACAAAG	CATTTTCAAC	AGATATGCTC	AGCAGGTATG	1740
	GTACCTAGCA	CCCAGGGGCC	TTACGTGGGA	TTGGAGAAAG	GGGACCTGAG	GGAGGGACAG	1800
	CCCTCACAGG	CCCTTACTGG	GATGCAGAGA	GGAGAAGTGA	CTTGATGGAC	TATTTTACCT	1860
	GCCTCTCTTC	CTGGATCGTC	TCCAGAACTG	CTGTGGCTGC	CAAGCTCGGT	AGAGACGTGG	1920
	CGCCCCACCC	AGTCTCATCC	GGGGGACTTC	AAGCTGGAAT	GCAGAGCTTA	GAAGGGGAGG	1980
15	GGATAATTAT	GGGGGTGTAG	GTGCATTGCC	CTCTAAATCT	TTAAACAAGC	AATTGGCAGT	2040
	ACCCCGTGAA	ACCTTTCTCT	CTCCTACTCG	GCCACCTCCC	ACCAACCTGG	CATCGTTCTT	2100
	CCCGGGAGCT	AGCCAGCTTC	AGAAAGCACA	TACAGCATCC	TTGCTGCCAA	ACCACCTATG	2160
	TGCACACAGG	ATTTCTCTAA	TGGCTTAATA	AACGTGTATA	AAGAACTCCT	TGACTTGTCA	2220
20	GAATAAAATA	GCTGCCAGGG	GCTCTGCACA	ATGAGCCTCT	TACCGTTAAA	AAAAAAAATA	2280
	AAAAAAAATA	AAAAAAAATA	AAAAAAAATA				

Seq ID NO: 327 Protein sequence:
Protein Accession #: BAB85075.1

25	1	11	21	31	41	51	
	MAYYQEPSVE	TSIIKFKDQD	FTTLRDHCLS	MGRTPKDETF	PAADSSIGQK	LLQEKRLSNV	60
	IWKRPQDLPG	GPPHFILDDI	SRFDIQGGGA	ADCWFLAALG	SLTQNPQYRQ	KILMVQSFSS	120
	QYAGIFRPFV	QCCQWVEVVF	IDDLRPVQGD	KCLFVRPRHQ	NOEFWPCLEL	KAYAKLLGSY	180
30	SDLHYGFLED	ALVDLTGGVI	TNIHLHSSPV	DLVKAVKTAT	KAGSLITCAT	PSGGTDTAQA	240
	MENGLVSLHA	YTVTGABQIQ	YRRGWEEIIS	LWNPWGNET	BWRGRWSDGS	QEWEECTCDPR	300
	KSLHLKRED	GEPWMSQDF	QOKFIAMPIC	SEIPITLDHG	NTLHEGWSQI	MFRKQVILGN	360
	TAGGPRNDQ	FNFSVQEPME	GTNVVVCVTV	AVTFSNLKAE	DAKFPLDFQV	ILAGSQKHCP	420
35	KLK						

Seq ID NO: 328 DNA sequence
Nucleic Acid Accession #: BC017490.1
Coding sequence: 74-2788

40	1	11	21	31	41	51	
	GTGGGTTCAG	TGAACCACTT	TTGCGCGGAA	ACCTGGTGTG	TGCTGTAGTG	GCGGAGAGGA	60
	TCGTGGTACT	GCTATGCGGG	AATCATCGGA	ATCCTTCAAC	ATGGCATCCA	GCCCGGCCCA	120
45	GGGTGCGCGA	GGCAATGATC	CTCTCACCTC	CAGCCCTGGC	CGAAGCTCCC	GGCGTACTGA	180
	TGCCCTCAC	TCCAGCCCTG	GCGGTGACCT	TCCACCATTT	GAGGATGAGT	CCGAGGGGCT	240
	CCTAGGCACA	GAGGGGCCCC	TGGAGGAAGA	AGAGGATGGA	GAGGAGCTCA	TTGAGATGG	300
	CATGGAAAGG	GACTACCGCG	CCATCCAGAG	GCTGGAGGCC	TATGAGGCGG	AGGAGCTGGC	360
	TCTGGATGAT	GAGGACGTAG	AGGAGCTGAC	GGCCAGTCAG	AGGAGGGCAG	CAGAGCGGGC	420
50	CATGCGGCAG	CGTGACCGGG	AGGCTGGCGG	GGGCTGGGCG	CGCATGCGCC	GTGGGCTCCT	480
	GTATGACAGC	GATGAGGAGG	ACGAGGAGCG	CCCTGCCCGC	AAGCGCCGCC	AGGTGGAGCG	540
	GGCCACGAGG	GACGGCGAGG	AGGACGAGGA	GATGATCGAG	AGCATCGAGA	ACCTGGAGGA	600
	TCTCAAAGGC	CACCTCTGTC	GCGAGTGGGT	GAGCATGGCG	GGCCCCGGCG	TGGAGATCCA	660
	CCACCGCTTC	AAGAAGCTTC	TGCGCACTCA	CGTCGACAGC	CACGCGCCCA	ACGTCTTCAA	720
55	GGAGCGCATC	AGCGACATGT	GCAAGAGAGAA	CGGTGAGAGC	CTGGTGGTGA	ACTATGAGGA	780
	CTTGGCAGCC	AGGGAGCAGC	TGCTGGCCTA	CTTCTCGCCT	GAGGCACCGG	CGGAGCTGCT	840
	GCAGATCTTT	GATGAGGCTG	CCCTGGAGGT	GGTACTGGCC	ATGTACCCCA	AGTACGACCG	900
	CATCACCAAC	CACATCCATG	TCCGCATCTC	CCAACCTGCT	CTGGTGGAGG	AGCTGCGCTC	960
	GCTGAGGAGC	CTGCATCTGA	ACCAGCTGAT	CGGCACCACT	GGGGTGGTGA	CCAGCTGCAC	1020
60	TGGCGTCTCG	CCCCAGCTCA	GCAATGGTCAA	GTACAACTGC	AACAAAGTGA	ATTTGCTCCT	1080
	GGGTCTCTTC	TGCCAGTCCC	AGAAACAGGA	GGTGAACCA	GGCTCCTGTC	CTGAGTGCCA	1140
	GTGCGCCGCG	CCCTTTGAGG	TCAACATGGA	GGAGACCATC	TATCAGAACT	ACCAGCGTAT	1200
	CCGAATCCAG	GAGAGTCCAG	GCAAAAGTGGC	GGCTGGCCGG	CTGCCCGGCT	CCAAGGACGC	1260
	CATTCTCCTC	GCAGATCTGG	TGGACAGCTG	CAAGCCAGGA	GACGAGATAG	AGCTGACTGG	1320
65	CATCTATCAC	AACAACATAT	ATGGCTCCCT	CAACACTGCC	AATGGCTTCC	CTGTCTTTGC	1380
	CACCTGTATC	CTAGCCCAAC	ACGTGGCCAA	GAAGGACAAC	AAGGTGCTGT	TAGGGGAAC	1440
	GACCGATGAA	GATGTGAAGA	TGATCACTAG	CCTCTCCAAG	GATCAGCAGA	TGGAGAGGAA	1500
	GATCTTTGCC	AGCATGCTCT	CTTCCATCTA	TGGTCATGAA	GACATCAAGA	GAGGCCTGGC	1560
70	TCTGGCCCTG	TTGCGAGGGG	AGCCCAAAAA	CCAGGTGGGC	AAGCACAAGG	TACGTGGTGA	1620
	TATCAACGTG	CTCTTGTGCG	GAGACCCCTG	CACAGCGAAG	TGGCAGTTTC	TCAAGTATAT	1680
	TGAGAAAGTG	TCCAGCCGAG	CCATCTTCAC	CACCTGGCCG	GGGGCGTCCG	CTGTGGGCTC	1740
	CACGGCGTAT	GTCAGCGGGC	ACCCTGTCCG	CAGGGAGTGG	ACCTTGGAGG	CTGGGGCCCT	1800
	GGTCTGCGCT	GACCGAGGAG	TGTGTCTCAT	TGATGAATTT	GACAAAGATGA	ATGACCAGGA	1860
	CAGAACCCAG	ATCCATGAGG	CCATGGAGCA	ACAGAGCATC	TCCATCTCGA	AGGCTGGCAT	1920
75	CGTCACTCC	CTGCAGGCTC	GCTGCACGGT	CATTGCTGCC	GCCAAACCCA	TAGGAGGGCG	1980
	CTAGCACCOC	TGCGTAGCTT	CTCTGTAGAA	CGTGGACCTC	ACAGAGCCCA	TCACTCTAAG	2040
	CTTTGACATC	CTGTGTGTGG	TGAGGGACAC	CGTGGACCCA	GTCCAGGACG	AGATGCTGGC	2100
	CCGCTTCGTG	GTGGGCAGCC	ACGTGAGACA	CCACCCAGC	AACAAGGAGG	AGCAGGGGCT	2160
	GGCCAAATGG	AGCGCTGCTG	AGCCCGCCAT	GCCCAACAGC	TATGGCGTGG	AGCCCTTGCC	2220
80	CCAGGAGGTC	CTGAGAGAGT	ACATCATCTA	CGCCAAAGGAG	AGGCTCCACC	CGAAGCTCAA	2280
	CCAGATGGAC	CAGGACAAGG	TGGCCAAAGT	GTACAGTGAC	CTGAGGAAAG	AATCTATGGC	2340
	GACAGGCAGC	ATCCCAATTA	CGGTGCGGCA	CATCGAGTCC	ATGATCCGCA	TGGCGGAGGC	2400
	CCAGCGCGCG	ATCCATCTGC	GGGACTATGT	GATCGAAGAC	GACGTCAACA	TGGCCATCCG	2460
	CGTGATGCTG	GAGAGCTTCA	TAGACACACA	GAAGTTTCAGC	GTGATGCGCA	GCATGCGCAA	2520
85	GACTTTTGCC	CGCTACCTTT	CATTCGGGCG	TGACAAACAAT	GAGCTGTGTC	TCTTCATACT	2580
	GAAGCAGTTA	GTGGCAGAGC	AGGTGACATA	TCAGCGCAAC	CGCTTTGGGG	CCAGCAGGGA	2640
	CACATTTGAG	GTCCCTGAGA	AGGACTTGGT	GGATAAGGCT	CGTCAGATCA	ACATCCACAA	2700

CCTCTCTGCA TTTTATGACA GTGAGCTCTT CAGGATGAAC AAGTTCAGCC ACGACCTGAA 2760
 AAGGAAAATG ATCTCTGCAGC AGTTCTGAGG CCTATGCCA TCCATAAGGA TTCCTTGGGA 2820
 TTCTGGTTTG GGGTGGTTCAG TGCCCTCTGT GCTTTATGGA CAAAAACCA GAGCACTTGA 2880
 TGAACCTCGG GTACTAGGGT CAGGGCTTAT AGCAGGATGT CTGGCTGCAC CTGGCATGAC 2940
 TGTTTGTTC TCCAAGCCTG CTTTGTGCTT CTCACCTTTG GGTGGGATGC CTGGCCAGTG 3000
 TGTCTTACTT GGTGTCTGAA CATCTTGCCA CTCCGAGTG CTTTGTCTCC ACTCAGTACC 3060
 TTGGATCAGA GCTGCTGAGT TCAGGATGCC TGGGTGTGTT TTAGGTGTTA GCCTTCTTAC 3120
 ATGGATGTCA GGAGAGCTGC TGCCCTCTTG GCGTGAGTTG CGTATTCAGG CTGCTTTTGC 3180
 TGCCCTTGGC CAGAGAGCTG GTTGAAGATG TTTGTAATCG TTTTCAGTCT CCTGCAGGTT 3240
 TCTGTGCCCC TGTGGTGGAA GAGGGCAGCA CAGTGCCAGC GCAGCGTTCT GGGCTCCTCA 3300
 GTCGACGGGG TGGGATGTGA GTCATGCGGA TTATCCACTC GCCACAGTTA TCAGCTGCCA 3360
 TTGCTCCCTG TCTGTTTCCC CACTCTCTTA TTTGTGCATT CGGTTTGGTT TCTGTAGTTT 3420
 TAATTTTTAA TAAAGTTGAA TAAATATAA AAAAAAAAAA AAAAAA

15 Seq ID NO: 329 Protein sequence:
 Protein Accession #: AAH17490.1

1 11 21 31 41 51
 20 MAESSESFTM ASSPAQRRRG NDPLTSSPGR SSRRTDALTS SPGRDLPPFE DSEGLLGTE 60
 GPLEEEEDGE ELIGDGMERD YRAIPELDAY EAEGLALDDE DVEELTASQR EAAERAMRQR 120
 DREAGRLAGR MRGLLYDSD EEDERPARK RRQVERATED GEEDEEMIES IENLEDLKGH 180
 SVREWVSMAG PRLEIHRPFK NFLRTHVDSH GHNVFKERIS DMCKENRESL VVNYEDLAAR 240
 EHVLAFLPE APABELQIPD EAALVVLAM YPKYDRITNH IHVRI SHLPL VEELRSLRQL 300
 25 HLNLQIRTSQ VVTSTCVLP QLSMVKYNCN KCNFVLGPFC QSONQEVKPG SCPEQCSAGP 360
 FEVNMETIY QNYQIRIQE SPGKVAAGRL PRSKDAILLA DLVDSCKPGD EIELTGIYHN 420
 NYDGLSLANT GPPVFPATVIL ANHVAKKDNK VAVGELTDED VKMITSLSKD QQIGEKIFAS 480
 TAPSIYGHED IKRGLALALF GGEPKPNPGGK HKVRGDIINV LCGDPGTAKS QFLKYIEKVS 540
 30 SRAIFTTGGQ ASAVGLTAYV QRHPVSREW L EAGALVLAD RGVCLIDEFD KMDQDRTSI 600
 HEAMEQQSIS ISKAGIVTSL QARCTVIAAA NPIGGRYDPS LTFSENVDLT EPIISRPDIL 660
 CVVRDVTDPV QDEMLARFVV GSHVRHHPSN KEEGLANGS AAEPAAMPNTY GVEPLPQEV 720
 KKYIIYAKER DKVAKMYSDL RKESMATGSI PITVRHIESM IRMAEAHARI 780
 HLRDVIIEED VNMAIRVMLE SFIDTQKFSV MRSMRKTFAR YLSFRDNDNE LLLFILKQLV 840
 35 AEQVTYQRNR PGAQDITIEV PEKDLVDRAR QINIHNLSAF YDSELFMRNK FSHDLKRKMI 900
 LQQF

40 Seq ID NO: 330 DNA sequence
 Nucleic Acid Accession #: M17254
 Coding sequence: 257-1645

1 11 21 31 41 51
 45 GTCCGCGCGT GTCCGCGCCC GCGTGTGCCA GCGCGCGTGC CTTCGCGGTG CGCGCCGAGC 60
 CGGGTGCAC TAACTCCCTC GCGCGCGACG GCGCGGCTAA CCTCTCGGTT ATTCAGGAT 120
 CTTTGGAGAC CCGAGGAAAG CCGTGTGAC CAAAAGCAAG ACAAATGACT CACAGAGAAA 180
 AAAGATGGCA GAACCAAGGG CAACTAAAGC CGTCAGGTTT TGAACAGCTG GTAGATGGGC 240
 TGGCTTACTG AAGGACATGA TTCAGACTGT CCGGACCCCA GCAGCTCATA TCAAGGAAGC 300
 CTTATCAGTT GTGAGTGAGG ACCAGTCTGT GTTTGAGTGT GCCTACGGAA CGCCACACCT 360
 50 GGCTAAGACA GAGATGACCG CGTCTCTCTC CAGCGACTAT GGACAGACTT CCAAGATGAG 420
 CCCACGGCTC CTTCCAGCAG ATTGGCTGTC TCAACCCCCA GCCAGGGTCA CCATCAAAAT 480
 GGAATGTAACT CCTAGCCAGG TGAATGGCTC AAGGAATCTT CCTGATGAAT GCAGTGTGGC 540
 CAAAGGCGGG AAGATGTGGT GCAGCCGAGA CACCGTTGGG ATGAATAACG GCAGCTACAT 600
 GGAGGAGAAG CACATGCCAC CCCCACCAT GACCACGAAC GAGCGCAGAG TTATCGTGCC 660
 55 AGCAGATCCT CTCTAGTGA GTACAGACCA TGTGCGGAG TGGCTGAGT GGGCGGTGAA 720
 AGAATATGGC CTCTCCAGAC TCAACATCTT GTTATTCAG AACATCGATG GGAAGGAAC 780
 GTGCAAGATG ACCAAGGACG ACTTCCAGAG GCTCACCACC AGCTACAACG CCGACATCCT 840
 TCTCTCACAT TCCCATACC TCAGAGAGAC TCCTCTTCCA CATTTGACTT CAGATGATGT 900
 TGATAAAGCC TTACAAAAC CTCCACGGTT AATGCATGCT AGAAACACAG ATTTACCATA 960
 60 TGAGCCCCC AGGAGATCAG CCTGGACGGG TCACGGCCAC CCCACGCCCC AGTCGAAAGC 1020
 TGCTCAACCA TCTCCTTCCA CAGTGCCCAA AACTGAAGAC CAGGTCTCTC AGTTAGATCC 1080
 TTATCAGATT CTGGACCAA CAAGTAGCCG CCTTGCAAA CAGGCGAGTG GCCAGATCCA 1140
 GCTTTGGCAG TTCCTCTCTG AGCTCTGTGC GGACAGCTCC AACTCCAGCT GCATCACCTG 1200
 GGAAGGCACC AACGGGGAGT TCAAGATGAC GGATCCCGAC GAGGTGGCCC GGGCTGGGG 1260
 65 AGAGCGGAAG AGCAAAACCA ACATGAAC TAATAGCTC AGCGCGCCCC TCCGTTACTA 1320
 CTATGACAAG AACATCATGA CCAAGGTCCA TGGGAAGGCG TACGCTTACA AGTTGACTT 1380
 CCACGGGATC GCCCAGGCCC TCACGCCCCA CCCCCCGGAG TCATCTCTGT ACAAGTACCC 1440
 CTCAGACCTC CCGTACATGG GCTCCTATCA CGCCCAACCA CAGAAGATGA ACTTTGTGGC 1500
 GCCCCACCCT CCAGCCCTCC CCGTGACATC TTCCAGTTTT TTTGCTGCCC CAACCCATA 1560
 70 CTGGAATTCA CCAACTGGGG GTATATACCC CAACACTAGG CTCCCCCA GGCATATGCC 1620
 TTCTCATCTG GGCACCTTACT ACTAAAGACC TGGCGGAGGC TTTTCCATC AGCGTGCACT 1680
 CACCAGCCCA TCGCCACAAA CTCTATCGGA GAACATGAAT CAAAAGTGCC TCAAGAGGAA 1740
 TGAATAAAGC TTTACTGGGG CTGGGGAAGG AAGCGGGGA AGAGATCCAA AGACTCTTGG 1800
 GAGGGAGTTA CTGAAGTCTT ACTACAGAAA TGAGGAGGAT GCTAAAATG TCAAGAAAT 1860
 75 GGACATATCA TCTGTGACT GACCTTGTA AAGACAGTGT ATGTAGAAGC ATGAAGTCTT 1920
 AAGGACAAAG TGCCAAAGAA AGTGGTCTTA AGAAATGTAT AAACCTTAGA GTAGAGTTTG 1980
 AATCCCACTA ATGCAAACTG GGATGAAACT AAAGCAATAG AAACAACACA GTTTTGACCT 2040
 AACATACCGT TTATAATGCC ATTTTAAGGA AAACCTACCTG TATTTAAAAA TAGTTTCATA 2100
 TCAAAAACAA GAGAAAAGAC ACGAGAGAGA CTGTGGCCCA TCAACAGACG TTGATATGCA 2160
 80 ACTGCATGGC ATGTGCTGTT TTGGTTGAAA TCAAAATACAT TCCGTTGAT GGACAGCTGT 2220
 CAGCTTTCTC AAACCTTGAA GATGACCCAA AGTTTCCAA CTTCTTACAG TATTACCGGG 2280
 ACTATGAAT AAAAGGTGGG ACTGAGGATG TGTATAGAGT GAGCGTGTGA TTGTAGACAG 2340
 AGGGGTGAAG AAGGAGGAGG AAGAGGCAGA GAAGGAGGAG ACCAGGCTGG GAAAGAACT 2400
 TCTCAAGCAA TGAAGACTGG ACTCAGGACA TTTGGGACT GTGTACAATG AGTTATGGAG 2460
 85 ACTCGAGGGT TCATGCAGTC AGTGTATAC CAAACCCAGT GTTAGGAGAA AGGACACAGC 2520
 GTAATGGAGA AAGGGAAGTA GTAGAATTCA GAAACAAAAA TGGCATCTC TTTCTTTGTT 2580
 TGTCAAATGA AAATTTTAAC TGAATTTGTC TGATATTTAA GAGAAACATT CAGGACCTCA 2640
 TCATTATGTG GGGGCTTTGT TCTCCACAGG GTCAGGTAAG AGATGGCCTT CTTCGCTGCC 2700

5
 10
 15
 20
 25

```

ACAATCAGAA ATCAGCGAGG CATTTTGGGT AGGCGGCTTC CAGTTTTCCT TTGAGTCGG 2760
AACGCTGTGC GTTTGTGAGA ATGAAGTATA CAAGTCAATG TTTTCCCCC TTTTATATA 2820
ATAATTATAT AACTTATGCA TTTATACACT ACGAGTTGAT CTCGGCCAGC CAAAGACACA 2880
CGACAAAAGA GACAAATCGAT ATAATGTGGC CTTGAATTTT AACTCTGTAT GCTTAATGTT 2940
TACAATATGA AGTTATTAGT TCTTAGAATG CAGAATGTAT GTAATAAAAT AAGCTTGGCC 3000
TAGCATGGCA AATCAGATTT ATACAGGAGT CTGCATTGTC ACTTTTTTTA GTGACTAAAG 3060
TTGCTTAATG AAAACATGTG CTGAATGTTG TGGATTTTGT GTTATAATTT ACTTTGTCCA 3120
GGAACCTGTG CAAGGGAGAG CCAAGGAAAT AGGATGTTTG GCACCC

```

Seq ID NO: 331 Protein sequence
 Protein Accession #: AA52398

15
 20
 25

```

1 11 21 31 41 51
MIQTVDPDAA HIKEALSUVS EDQSLFECAY GTPHLAKTEM TASSSSDYGQ TSKMSPRVPQ 60
QDWLSQPPAR VTIKMECNPS QVNGSRNSPD ECSVAKGGKM VGSPTDVGMM YGSYMBEEKHM 120
PPFNMTTNER RVIVPADPTL WSTDHVRQWL EWAVKEYGLP DVNILLFQNI DGKELCKMTK 180
DDFQRLTSPY NADILLSHLH YLRETPHLPHL TSDDVDKALQ NSPRLMHARN TDLPEYPPRR 240
SAWTGHGHPT PQSKAAQPSF STVPKTEDQR PQLDPYQILG PLSRLANPG SQQLQLWQFL 300
LELLSDSSNS SCITWEGTNG EPKMTDPDEV ARRWGERKSK PNMNYDKLSR ALRYYYDKNI 360
MTKVHGKRYA YKDFPHGIAQ ALQPHPPSS LYKYPSDLPY MGSYHAHPQK MNFVAPHPPA 420
LPVTSSTFFA APNPNVNSPT GGIYPNTRLP TSHMPSHLGT YY 462

```

Seq ID NO: 332 DNA sequence
 Nucleic Acid Accession #: NM_000020
 Coding sequence: 283-1794

30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

```

1 11 21 31 41 51
AGGAAACGGT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA 60
AGAAACATTT TTGCTCCAGC CCCCATCCCA GTCCCGGGAG GCTGCCCGGC CAGCTGCGCC 120
GAGCGAGCCC CTCGCCCGGT CCAGCCCGGT CCGGGGCGCG GCGGACCCCG AGCCCGCGGT 180
CCAGCGCTGG CGGTGCAACT GCGGCCGCGC GGTGGAGGGG AGGTGGCCCC GGTCCGCGCA 240
AGGCTAGCGC CCGGCCACCC GCAGAGCGGG CCCAGAGGGA CCAATGACCTT GGGCTCCCCC 300
AGGAAAGGCC TTCTGATGCT GCTGATGGCC TTGGTGACCC AGGGAGACCC TGTGAAGCG 360
TCTCGGGGCC CGCTGGTGAC CTGCACGTGT GAGAGCCAC ATTGCAAGGG GCCTACCTGC 420
CGGGGGGCGT GGTGCACAGT AGTCTGTGTG CCGGAGGAGG GGAGGCACCC CCAGGAACAT 480
CGGGGCTGCG GGAACCTTGA CAGGGAGCTC TGCAGGGGCG GCCCCACCGA GTTCTGCAAC 540
CACTACTGCT GCGACAGCCA CCTCTGCAAC CACAACGTGT CCTGTGTGCT GGAGGCCACC 600
CAACCTTCCTT CGGAGCAGCC GGGAAACAGAT GGCCAGCTGG CCTGTATCCT GGGCCCGGTG 660
CTGGCCTTGC TGGCCTCGGT GGCCTGGGCT GTCTGGGCGT TGTGGCATGT CCGACGGAGG 720
CAGGAGAGAG AGCGTGGGCT GCACAGCGAG CTGGGAGAGT CCACTCTCAT CCTGAAGCA 780
TCTGAGCAGG GCGACACGAT GTTGGGGGAC CTCTTGACAC GTGACTGCAC CACAGGGAGT 840
GGCTCAGGGC TCCCTTCTCT GGTGCAGAGG ACAGTGGCAC GGCAGGTGTC CTTGTGGAG 900
TGTGTGGGAA AAGGCGGCTA TGGCGAAGTG TGGCGGGGCT TGTGGCACGG TGAGAGTGTG 960
GCGCTCAGAA TCTTCTCTCT GAGGGATGAA CAGTCTCTGT TCGGGGAGAG TGAGATCTAT 1020
AACACAGTAT TGCTCAGACA GCACAAATC CTAGGCTTCA TCGGCTCAGA CATGACCTCC 1080
CGCAACTCGA GCACGCGAGT GTGGCTCATC ACGCACTACC ACGAGCACGG CTCCTCTTAC 1140
GACTTTCTGC AGAGACAGAC GCTGGAGCCC CATCTGGCTC TGAGGCTAGC TGTGTCCGG 1200
GCATGCGGCG TGGCGCACCT GCACGTGGAG ATCTTCGGTA CACAGGGCAA ACCAGCCATT 1260
GCCACCGCG ACTTCAAGAG CCGCAATGTG CTGCTCAAGA GCAACCTGCA GTGTGTGATC 1320
GCCGACCTGG GCTTGGCTGT GATGCACTCA CAGGGCAGCG ATTACCTGGA CATCGGCAAC 1380
AACCCGAGAG TGGGCACCAA GCGGTACATG GCACCCGAGG TGCTGGACGA GCAGATCCGC 1440
ACGACTGCTT TTGAGTCTTA CAAGTGGAGT GACATCTGGG CCTTGGGCTT GGTGTGTG 1500
GAGATTGGCC GCGGACCAAT CGTGAATGGC ATCGTGGAGG ACTATAGACC ACCCTTCTAT 1560
GATGTGGTGC CCAATGACCC CAGCTTTGAG GACATGAAGA AGTGGTGTG TGTGGATCAG 1620
CAGACCCCCA CCATCCCTAA CCGGCTGGCT GCAGACCCCG TCCTCTCAGG CCTAGCTCAG 1680
ATGATGCGGG AGTGTCTGGT ACCAAACCCC TCTGCCGAGC TCACCCGCTT GCGGATCAAG 1740
AAGACACTAC AAAAATTAG CAACAGTCCA GAGAAGCCTA AAGTGATTCA ATAGCCACAG 1800
AGCACCTGAT TCCTTTCTGC CTGCAGGGGG CTGGGGGGGT GGGGGGCAGT GGATGGTGG 1860
CTATCTGGGT AGAGGTAGTG TGAGTGTGGT GTGTGCTGGG GATGGGCAGC TGCGCCTGCC 1920
TGCTCGGCCC CCAGCCACCC CAGCCAAAAA TACAGCTGGG CTGAAACCTG ATCCCTGCTG 1980
GTCTGGCCTG CTCAAAGCGG CAGGCTCCTT GACGCTGGC TCTCTCCCA CCCCTATGGC 2040
CAGCATGGTG CACCCCTTAC CACTCCCGGG ACAGGATGCA AAGAGGCTC CAGAGTCAGA 2100
GTGCCAAGCC AGGGAATCCC AGTCCCAGAC TCAGAGCCCG GGCCTGCACT TTGCCCTG 2160
CCCTTGATCA ACCCACTGC CCCACAGAG CTGCCAGGGT GGCACAGGGC CCTGTCCAGC 2220
CCCTGGCACA CACTTCCCTG CCAGGCTCTA GCCTCTAGCA TAAGCTCCAG AGAGCCAGG 2280
CCCATCAGTT TCTCTCTGTG GATTGTATC TCAGCTCCAT GATGCTTGG GCTTTCTGTC 2340
TCCTCAACAA GAGTGAGCTT TGCTGAATGT CAGCTGCCTG AGAGAGCTGG GGCTGACTT 2400
ACTAGGGCAT TAAATCTTAA GAGTCTCTAC TGAGGTGTGG CAGGATCACA GGCCAGTGGA 2460
AAAAGGCGAG GTCAGATGGG CAAGGCCAGG GACTTTCAGA TTAAGTGAGA GGATATGAG 2520
GCCAAGCATG GCAGGGGGAA GGTCACTGGG TGTCAAGAGA CCCAGGCTG ACCCGGATG 2580
TTTGCTCCAT GTGACAAAAG CAGGCTCTGC TCAGGACCTT TTCTTTTCTT TTTTCTTCT 2640
TTTTTTTTTT GACACGGAGT TTCGCTCTTG TTGTCCAGGC TAGAGTGCAA TGGCATGATC 2700
CCAGCTCACC GCAACGTCTA CCTCCAGGT TCAAAATCAT CTCTTGCTTC AGACTCCCGA 2760
GTAGCTGGGA TTACAGGCAC ATGCCACCAT GCCTGGCTAA TTTTGTATAT TTAGTAGAAA 2820
CAGGGTTTCA CCATGCTGGC CATGCTGGTT CTCGAACTCC TGACCTCAGG TGTTCACCT 2880
ACCTCAGGCT GCCAAAGTGG TGGGGTTACA GGTGTGAGCC ATCGGCGCTG GCCAGGACCT 2940
TTGTTTCTTA TCTACATATT GGAAGATTGG GTCTCTGATG CCTTGAAGGC TTCTTATGCT 3000
CTAGTCTCTG GACACTCTAG CCTATATCAC AGCTAACTTC YTCAGTCTCA TCTATTCTTT 3060
ATGCTCCAGC CCCTGGCAAT TTGCTTCAAG ATGGGGGTTT GAAATAACTT TTACCTGACT 3120
CAAGGAGTGT CTGGAGCACC TCCTAGTCTA AGTCTGCAAG CTCAGTCTCT TGCTTAAAC 3180
CATGGCAGTG GCCACCTTGG GGCTCAGACA GCTCTGGGCC TTTTGACCA AAGCCAGCCC 3240
CTGCGCCTCT CTGTGGCATA GTCTTCTCTG CCCAGGAGT GCAGGGCGGC TTCTTCAAG 3300
GCTTCCAGG CTCAAGAAAT ATTTGGCTCC ATCCAAGAG GCTCCAGCTC CCCTACTGGC 3360
CCCTGGCTTC AGGCCACAC CCCTGGGCCA GGCACAGAGA GTGTGTCTCA GGAGAAATCA 3420
ATGGGCTCTA GAGAGACACA CAGAAAGTTT GGGCATTGGG GAAATTTTCA AGGRTGTATG 3480

```

5 TATGGYTCAC GTATGGWGCA GGTGTGCTCTG GTCCYKGGGT GCAGGGAAGT GGGCTGCAGG 3540
 GAAGTGGATT GGAGGGGAGC TTGAGGAATA TAAGGAGCGG GGGTGGAGAC TCAGGCTATG 3600
 GACAAGGACA GCCCCAGAGT TGGGAAGACC TGGCCTTAGT CGTCTCAGC CTAGGGCAGG 3660
 GCAGTGAACA AAGCTCTCCC CGCTCCTGCT GTAATGACCC AGAGTAGCCT CCCAGGCCG 3720
 GCATCTTATG TGTGTCTTCC ACCATCTCTCA TGGTGGCACT TTTCTAGGCC TGTCTCCAG 3780
 CATTTGGCAA GGCTCGGAAG AGAACCAGGA AGTGAACCTG GGTGAAAACA GAAAGCTCAA 3840
 TGGATGGGCT AGGTTCCAG ATCATTAGGG CAGAGTTTGC ACGTCTCTG GTTCACTGGG 3900
 AATCCACCA GCCCAGGAAT CATCTCCCTC TTTGAAGGAT TTTWATTCT ACTGGGTTTT 3960
 10 GGAACAACT CCTGCTGAGA CCCACAGCC AGAACTGAA AGCAGCAGCT CCCCAAAGCC 4020
 TGGAAAATCC CTAAGAGAAG GCCTGGGGGA MAGGAATGG AGTGACAGGG GACAGGTAGA 4080
 GAGAAGGGGG CCCAATGGCC AGGGAGTGAA GGAGGTGGCG TTGCTGAGAG CAGTCTGCAC 4140
 ATGCTTCTGT CTGAGTGCAG GAAGGTGTTC CAGGGTCGAA ATTACACTTC TGTACCTGG 4200
 AGACGCTGTT TGTGGGAGCA CTGGGCTCAT GCCTGGCACA CAATAGGTCT GCAATAAACC 4260
 15 ATGGTTAAAT CCTGAAAAA AAAAAAAA

Seq ID NO: 333 Protein sequence
 Protein Accession #: NP_000011

20 1 11 21 31 41 51
 MTLGSPRKGL LMLLMALVTQ GDPVKPSRGP LVTCTCESPH CKGPTCRGAW CTVVLVREEG 60
 RHPQEHRCG NLHRELRCRG PTEFVNHYCC DSHLCNHNVS LVLEATQPPS EQPGTDGQLA 120
 25 LILGVLALL ALVALGVGL WHVRRRQEKQ RGLHSELGES SLILKASEQG DTMLGDLDS 180
 DCTTGGSSGL PFLVQRTVAR QVALVECVGK GRYGEVWRGL WHGESVAVKI FSSRDEQSWF 240
 RETEIYNTVL LRHDNILGFI ASDMTSRNSS TQLWLITHYH EHGSLYDFLQ RQTLLEPHLAL 300
 RLAVSAACGL AHLHVEIFGT QGKPAIAHRD FKSRNVLVKS NLQCCIALDLG LAVMHSQSGD 360
 YLDIGNPRV GTKRYMAPEV LDEQIRTDCE ESYKNWDIWA FGLVLWEIAR RTIVNGIVED 420
 YRPPFYDVVF NDPSFEDMKK VVCVDQQTPT IPNRLAADPV LSGLAQMRE CWPFPNSARL 480
 30 TALRIKKTQL KISNSPEKPK VIQ

Seq ID NO: 334 DNA sequence
 Nucleic Acid Accession #: NM_004126.1
 Coding sequence: 108-329

35 1 11 21 31 41 51
 GGCACGAGCT CGTCCCGGCC TTCAGTTGTT TCGGGACGCG CCGAGCTTCG CCGCTCTTCC 60
 AGCGGCTCCG CTGCCAGAGC TAGCCCGAGC CCGGTTCTGG GCGGAAATG CCGTCCCTTC 120
 40 ACATCGAAGA TTTGCCAGAG AAGGAAAAAC TGAAATGGA AGTTGAGCAG CTTCCCAAAG 180
 AAGTGAAGTT GCAGAGACAA CAAGTGTCTA AATGTTCTGA AGAAATAAG AACTATATTG 240
 AAGAACCTTC TGSAGAGGAT CCTCTAGTAA AGGGAATTCC AGAAGACAAG AACCCCTTTA 300
 AAGAAAAAGG CAGCTGTGTT ATTTCAATAA TAACCTGGGA GAACTGCAT CTAAGTGGGA 360
 AGAAGTAGTT TGTTTTAGTT TTCCAGATA AAACCAACAT GCTTTTAAAG GAAGGAAGAA 420
 45 TGAATTAAG AGGAGACTTT CTTAAGCACC ATATAGATAG GGTATGTAT AAAAGCATAT 480
 GTGCTACTCA TCTTTGCTCA CTATGCAGTC TTTTAAAGA GAGCAGAGAG TATCAGATGT 540
 ACAATTATGG AATAAGAACC ATTACTTGAG CATGACACTT CTTTCAGTAT ATTGCTTGAT 600
 GCTTCAATA AAGTTTGTG TT

Seq ID NO: 335 Protein sequence
 Protein Accession #: NP_004117.1

50 1 11 21 31 41 51
 55 MPALHIEDLP EKEKLEMEVE QLRKEVKLQR QVSKCSEBI KNYIEERSGE DPLVKGIPED 60
 KNPFKEKGSC VIS

Seq ID NO: 336 DNA sequence
 Nucleic Acid Accession #: NM_005795
 Coding sequence: 555-1940

60 1 11 21 31 41 51
 65 GCACGAGGGA ACAACCTCTC TCTCTSCAGC AGAGAGTGTC ACCTCCTGCT TTAGGACCAT 60
 CRAAGCTGCG TAAGTAATC TCATCCTAAT TGCAGGATCA CATTGCAGAG CTTTCACTCT 120
 TTCCCACTTT GCTGTGGGTT AAATCTCTTC TGGGGAATCT CAGAAAGTAA AGTTCCATCC 180
 TGAGAAATTT TCACAAAGAA TTTCTTAAAG AGCTGGAGCTG GGTCTTGACC CCGGAATTT 240
 70 AAGAAATTTT TAAAGACAA GTCAATATG ATCCAAGAGA AAATGTGATT TGAGTCTGGA 300
 GACAAATGCT CATATCGTCT AATAATAAAA ACCCATACTA GCCTATAGAA AACAATATT 360
 GAATAATAAA AACCCATCT AGCCTATAGA AAACAATATT TGAAGATTG CTACCACTAA 420
 AAAGAAACT ACTACAACCT GACAAGACTG CTGCAAACTT CAATTGGTCA CCACAACCTG 480
 ACAAGGTTTC TATAAAACAA GATTGCTACA ACTTCTAGTT TATGTTATAC AGCATATTTT 540
 75 ATTTGGGCTT AATGATGGAG AAAAAGTGTA CCCTGTATT TCTGGTCTCT TTGCCTTTT 600
 TTATGATTCT TGTACAGCA GAATTAGAAG AGAGTCCCTGA GGACTCAATT CAGTTGGGAG 660
 TTACTAGAAA TAAATCATG ACAGCTCAAT ATGAATGTTA CCAAAAGATT ATGCAAGACC 720
 CCATTACACA AGCAGAGGCG GTTACTGCA ACAGAACTTG GATGGATGG CTCTGCTGGA 780
 AGAATGTTGC AGCAGGAAC TCAATGATGC AGCTCTGCCC TGATTACTTT CAGGACTTTG 840
 80 ATCCATCAGA AAAAGTTACA AAGATCTGTG ACCAAGATGG AAAGTGGTTT AGACATCCAG 900
 CAGACACAG ACATGAGCA AATTATACCC AGTGAATGT TAACACCCAC GAGAAAGTGA 960
 AGACTGCACT AAATTTGTTT TACCTGACCA TAATTGGACA CGGATTGTCT ATTGCATCAC 1020
 TGCTTATCTC GCTTGGCATA TTCTTTTATT TCAAGAGCCT AAGTTGCCAA AGGATTACCT 1080
 TACACAAAAT TCTGTTCTTC TCATTGTTT GTAACCTGT TGAACAATC ATTCACTCA 1140
 CTGCAAGTGGC CACAAACAG GCCTTAGTAG CCACAAATCC TGTATGTGCG AAAGTGTCCC 1200
 85 AGTTCATTCA TCTTTACCTG ATGGGCTGTA ATTACTTTTG GATGCTCTGT GAAGGCATT 1260
 ACCTACACAC ACTCATTTGT GTGGCGGTGT TTGCAGAGAA GCAACATTTA ATGTGGTATT 1320
 ATTTCTTGG CTGGGGATT CCACGTATTC CTGCTGTAT ACATGCCATT GCTAGAAGCT 1380

TATATTACAA TGACAAATTGC TGGATCAGTT CTGATACCCA TCTCCTCTAC ATTATCCATG 1440
 GCCCAATTTG TGCTGCTTTA CTGGTGAATC TTTTTCCTT GTTAAATATT GTACGCGTTC 1500
 TCATCACCAA GTTAAAGATT ACACACCAAG CGGAATCCAA TCTGTACATG AAAGCTGTGA 1560
 GAGCTACTCT TATCTTGGTG CCATTGCTTG GCATTGAATT TGTGCTGATT CCATGGCGAC 1620
 5 CTGAAGSAAA GATTGCGAGG GAGGTATATG ACTACATCAT GCACATCCTT ATGCACTTCC 1680
 AGGGTCTTTT GGTCTCTACC ATTTCTGCT TCTTTAATGG AGAGGTTCAG GCAATTCTGA 1740
 GAAGAACTG GAATCAATAC AAAATCCAAT TTGGAACAG CTTTCCAAC TCAGAGCTC 1800
 TTGTAGTGC GTCTTACACA GTGTCAACAA TCAGTATGAG TCCAGGTAT AGTCATGACT 1860
 10 GTCTAGTGA ACACCTTAAAT GGAAAAAGCA TCCATGATAT TGAAAAATGTT CTCTTAAAC 1920
 CAGAAAAATT ATATAATTGA AAATAGAAGG ATGGTTGTCT CACTGTTTGG TGCTTCTCCT 1980
 AACTCAAGGA CTTGACCCCA TGACTCTGTA GCCAGAAGAC TTCAATATTA AATGACTTTG 2040
 GGGAAATGCA TAAAGAAGAG CCTTCACATG AAATTAGTAG TGTGTTGATA AGAGTGTAAAC 2100
 ATCCAGCTCT ATGTGGGAAA AAAGAAATCC TGGTTTGTAA TGTTTGTGAG TAAATACTCC 2160
 CACTATGCTCT GATGTGACCG TACTAACCTG ACATCACCAA GTGTGGAATT GGAGAAAAGC 2220
 15 ACAATCAACT TTTCTGAGCT GGTGTAAGCC AGTTCACGCA CACCATTGAT GAATTCAAAC 2280
 AAATGCTGCT AAAACTTAAAC ATACATGTTG GGCATGATTC TACCCTTATT CCCCCAAGA 2340
 GACCTAGCTA AGGTCTATAA ACATGAAGGG AAAATTAGCT TTTAGTTTAA AAACCTCTTA 2400
 TCCCATCTTG ATTTGGGCGAG TTGACTTTTT TTTTTCCTCA GAGTCCGTA GTCCCTTTTG 2460
 TAACTACCTC CTCAATGGA CAATACCAGA AGTGAATTAT CCCTGCTGGC TTTCTTTTCT 2520
 20 CTATGAAAG CAACGAGTA CAATTGTTAT GATCTACTCA TTTGCTGACA CATCAGTTAT 2580
 ATCTTGTCG ATATCCATTG TGGAACTGG ATGAACAGGA TGTATAATAT GCAATCTTAC 2640
 TTCTATATCA TTAGGAAAAC ATCTTAGTTG ATGCTACAAA ACACCTTGTC AACCTCTTCC 2700
 TGTCTTACCA AACAGTGGGA GGGAAATCCT AGCTGTAAAT ATAAATTTTG CCCTTCCATT 2760
 TCTACTGTAT AAACAAATTA GCAATCATTT TATATAAGA AAATCAATGA AGGATTTCTT 2820
 25 ATTTTCTGG AATTTTGTAA AAAGAAATG TGAAAAATGA GCTTGTAAT ACTCCATTAT 2880
 TTTATTTTAT AGTCTCAAT CAAATACATA CAACCTATGT AATTTTAA GCAATATAT 2940
 AATGCAACAA TGTGTGTATG TTAATATCTG ATACTGTATC TGGGCTGATT TTTTAAATAA 3000
 AATAGAGTCT GGAATGCT

30 Seq ID NO: 337 protein sequence
 Protein Accession #: NP_005786.1

1 11 21 31 41 51
 35 MEKKCTLYFL VLLPFFMILV TAELEESPED SIQLGVTRNK IMTAQYECYQ KIMQDPIQQA 60
 EGVYCNRTWD GWLWNVDAVA GTESMQLCPD YFQDFPSEK VTKICDQDGN WFRHPASNRT 120
 WNTYTQCNVN THEKVTALN LEYLTIIHGG LSIASLLISL GIFFYFKSL S QORITLHKNL 180
 PFSFVCNSVV TIIHLTAVAN NQALVATNPV SCKVQPIHL YLNGCNVFWM LCEGIYLETL 240
 40 IVVAVPAERQ RLWNYFLWGF PLIPACIH AIARSLYND NCVISSDTHL LYIIRGPICA 300
 ALLVNLFFLL NIVRLITLKL KVTHQAESNL YMKAVRATLI LVPLLGIEFV LIPWRPEGKI 360
 AEEVYDIYIH ILMHFQGLLV STIFCFNGE VQAILRRNWN QYKIQFGNSF SNSEALRSAS 420
 YTVSTISDGP GYSHDCPSEH LNKSIHDIH NVLLKPENLY N

45 Seq ID NO: 338 DNA sequence
 Nucleic Acid Accession #: NM_001795
 Coding sequence: 25-2379

1 11 21 31 41 51
 50 GCAAGATCTG TTCTCTCTGG GAAGATGCAG AGGCTCATGA TGCTCCTCGC CACATCGGGC 60
 GCCTGCCTGG GCCTGCTGGC AGTGGCAGCA GTGGCAGCAG CAGGTGCTAA CCCTGCCCAA 120
 CGGGACACCC ACAGCCTGCT GCCACCCAC CGGCGCCAAA AGAGAGATTG GATTGGAAAC 180
 CAGATGCACA TTGATGAAGA GAAAAACACC TCACCTCCCC ATCATGTAGG CAAGATCAAG 240
 TCAAGCGTGA GTCCCAAGAA TGCCCAAGTAC CTGCTCAAAG GAGAATATGT GGGCAAGGTC 300
 55 TTCCGGGTGG ATGCAGAGAG AGGAGAGCGTG TTGCGCATTT AGAGGCTGGA CGGGAGAAAT 360
 ATCTCAGAGT ACCACCTCAC TGCTGTCTAT GTGGACAAAG ACACTGGTGA AAACCTGGAG 420
 ACTCCTTCCA GCTTCACCAT CAAAGTTCAT GACGTGAACG ACACTGGGCC TGTGTCTACG 480
 CATCGGTTGT TCAATGCGTC CGTGCTGAG TCGTGGCTG TGGGGACCTC AGTCATCTCT 540
 GTGACAGCAG TGGATGCAGA CGACCCCACT GTGGGAGACC ACGCTCTGT CATGTACCAA 600
 60 ATCTCGAAGG GGAAGAAGTA TTTTCCATC GATAATTCTG GACGTATTAT CACAATAACG 660
 AAAAGCTTGG ACCGAGAGAA GCAGGCCAGG TATGAGATCG TGGTGGAAAG CGGAGATGCC 720
 CAGGGCCTCC GGGGGGACTC GGGCACGGCC ACCGTGCTGG TCACTCTGCA AGACATCAAT 780
 GACAACTTCC CTTCTTCCAC CCAGACCAAG TACACATTG TCGTGCCTGA AGACACCCGT 840
 GTGGGCACTC CTGTGGGCTC TCTGTTGTTT GAGGACCCAG ATGAGCCCCA GAACCGGATG 900
 65 ACCAAGTACA GCATCTTGGC GGGCGACTAC CAGGACGCTT TCACCATGGA GACAAACCCC 960
 GCCACCAACG AGGCATCAT CAAGCCATG AAGCCTCTGG ATTATGAATA CATCCAGCAA 1020
 TACAGCTTCA TCGTCCAGGC CACAGACCCC ACCATCGACC TCCGATACAT GAGCCCTCCC 1080
 GCGGGAAACA GAGCCAGGT CATTATCAAC ATCACAGATG TGGACGAGCC CCCCATTTTC 1140
 CAGCAGCCTT TCTACCACT CCAGCTGAAG GAAAACAGA AGAAGCCTCT GATTGGCACA 1200
 70 GTGCTGGCCA TGGACCCCTGA TGGCGTAGG CATAGCATTG GATACTCCAT CCGCAGGACC 1260
 AGTGACAGG GCCAGTTCTT CCGAGTCACA AAAAAGGGGG ACATTACAA TGAGAAAGAA 1320
 CTGGACAGAG AAGTCTACCC CTGGTATAAC CTGACTGTGG AGGCCAAGA ACTGGATTCC 1380
 ACTGGAACCC CCACAGGAAA AGAATCCATT GTGCAAGTCC ACATTGAAGT TTTGGATGAG 1440
 AATGACAATG CCCCGAGTT TGCCAAAGCCC TACCAGCCCA AAGTGTGTGA GAACGCTGTC 1500
 75 CATGGCCAGC TGGTCTGCA GATCTCCGCA ATAGACAAGG ACATAACACC ACGAAACGTG 1560
 AAGTCAAAAT TCACCTTGAA TACTGAGAAC AACTTTACCC TCACGGATAA TCACGATAAC 1620
 ACGGCCAACA TCACAGTCAA GTATGGGCAG TTTGACCGGG AGCATAACCA GTTCCACTTC 1680
 CTACCCGTGG TCATCTCAGA CAATGGGATG CCAAGTCCGA CGGGCACCAG CACGCTGACC 1740
 GTGGCGGTGT GCAAGTCAA CGAGCAGGGC GAGTTCACTT TCTGCGAGGA TATGGCCGCC 1800
 80 CAGGTGGCGG TGGCATCCA GGCAGTGSTA GCCATCTTAC TCTGCTCTCT CACCATCACA 1860
 GTGATCACCC TGCTCATCTT CTTGCGGCGG CGGCTCCGGA AGCAGGCCCG CGCGCACGGC 1920
 AAGAGCGTGC CGGAGATCCA CGAGCAGCTG GTCACTACG ACGAGGAGGG CGGCGCGGAG 1980
 ATGGACACCA CTGCTACGCA TGTGTGGTGG CTCACCTCGG TGCGCGCGGG CGGGCCAAAG 2040
 CCCCCTGGCG CCGCGCTGGA CGCCCGGCTT TCCCTCTATG CGCAGGTGCA GAAGCCACCG 2100
 85 AGGCACCGCG CTGGGCGACA CGGAGGGCCC GGGGAGATGG CAGCCATGAT CGAGGTGAAG 2160
 AAGGACGAGG CGGACCAAGA CGGCGACGGC CCCCCTACG ACACGCTGCA CATCTACGGC 2220
 TACGAGGGCT CCGAGTCCAT AGCCGAGTCC CTCAGCTCCC TGGGCACCGA CTCATCCGAC 2280

	TCTGACGTGG	ATTACGACTT	CCTTAACGAC	TGGGGACCCA	GGTTTAAGAT	GCTGGCTGAG	2340
	CTGTACGCT	CGGACCCCG	GGAGGAGCTG	CTGTATTAGG	CGGCCGAGGT	CACCTCTGGC	2400
	CTGGGGACCC	AAACCCCTCG	CAGCCAGGCG	CAGTCAGACT	CCAGGCCACCA	CAGCCTCCAA	2460
5	AAATGGCAGT	GACTCCCGAG	CCGAGCACCC	CTTCTCGTG	GGTCCAGAG	ACCTCATCAG	2520
	CCTTGGGATA	GCAAACTCCA	GGTTCTGAA	ATATCCAGGA	ATATATGTCA	GTGATGACTA	2580
	TTCTCAAAATG	CTGGCAAAATC	CAGGCTGGTG	TTCTGTCTGG	GCTCAGACAT	CCACATAACC	2640
	CTGTCAACCA	CAGACCGCCG	TCTAACTCAA	AGACTTCCTC	TGGCTCCCA	AGGCTGCAAA	2700
	GCAAAACAGA	CTGTGTTTAA	CTGCTGCAGG	GTCTTTTCT	AGGCTCCCTG	AACGCTCTGG	2760
10	TANGCTGGT	GAGGTCTGG	TGCTATCTG	CCTGGAGGCA	AAGGCTTGA	CAGCTTGACT	2820
	TGTGGGGCAG	GATTCTCTGC	AGCCCATTC	CAAGGGAGAG	TGACCATCAT	GCCCTCTCTC	2880
	GGGAGCCCTA	GCCTCTCTCC	AACTCCATAC	TCCACTCCAA	GTGCCCCACC	ACTCCCCAAC	2940
	CCCTCTCCAG	GCCTGTCAA	AGGGAGGAAG	GGGCCCCATG	GCAGCTCCTG	ACCTTGGGTC	3000
	CTGAAGTGAC	CTCACTGGCC	TGCCATGCCA	GTAACTGTGC	TGTACTGAGC	ACTGAACACC	3060
15	ATTACAGGAA	ATGCTTATTA	AACTTGAAG	CAACTGTGAA	TTCACTCTGG	AGGGGCAGTG	3120
	GAGATCAGGA	GTGACAGATC	ACAGGGTGAG	GGCCACCTCC	ACACCCACCC	CCTCTGAGA	3180
	AGGCTTGAA	GAGCTGAGAC	CTTGCTTTGA	GACTCTCTAG	CACCCCTCCA	GTTTTCCTG	3240
	AGAAAGGCGA	GATGTTCCCG	GAGATCAGAA	GAGCTCTCCC	CTTCTCTGCC	TCACCTGGTC	3300
	GCCAAATCCAT	GCTCTCTTTC	TTTTCTCTGT	CTACTCTCTA	TCCCTTGGTT	TAGAGGAACC	3360
	CAAGATGTGG	CCTTTAGCAA	AACTGACAAT	GTCCAACCC	ACTCATGACT	GCATGACGGA	3420
20	GCCGAGCATG	TGTCTTTACA	CCTCGCTGTT	GTCACTCTC	AGGGAAGTGA	CCCTCAGGCA	3480
	CACCTTGAG	AAGGAAGGCC	CTGCCCTGCC	CAACCTCTGT	GGTCACCCAT	GCATCATTCC	3540
	ACTGGAAGCT	TTCTGTCAA	ACACACCTTG	GAGAAGTGGC	ATCAGTCAAC	AGAGAGGGGC	3600
	AGGGAAGGAG	ACACCAAGCT	CACCTTCGT	CATGGACCGA	GGTCCCACCT	CTGGCAAAGC	3660
	CCCTCACACT	GCAAGGGATT	GTAGATAACA	CTGACTTGT	TGTTTTAACC	AATAACTAGC	3720
25	TTCTTATAT	GATTTTTTTA	CTAATGATAC	TTACAAGTTT	CTAGCTCTCA	CAGACATATA	3780
	GAATAAGGGT	TTTTTGATAA	TAAGCAGGTT	GTTATTTAGG	TAAACATAT	TAATTCAGGT	3840
	TTTTTAGTTG	GAAAAACAAT	TCCTGTAAAC	TTCTATTTTC	TATAATTGTA	GTAATTGCTC	3900
	TACAGATAAT	GTCTATATAT	TGGCCAAACT	GGTGCATGAC	AAGTACTGTA	TTTTTTTATA	3960
30	CCTAATAATA	GAAAAATCTT	TAGCCTGGGC	AACAAAAAAA			

Seq ID NO: 339 Protein sequence
Protein Accession #: NP_001786

	1	11	21	31	41	51	
35	MQRIMMLLAT	SGACLGLLAV	AAVAAAGANP	AQRDTHSLLE	THRRQKRWDI	WNQMHIDEEK	60
	NTSLPHEHVK	IKSSSVSRKNA	KYLLKGEYVG	KVFRVDAETG	DVFAIERLDR	ENISEYHLTA	120
	VIVDKDTGEM	LETSSFTIK	VHDVNDNWPV	PTHRLFNASV	PESSAVGTSV	ISVTAVDADD	180
	PTVGDBASVM	YQILKGEYF	AIDNSGRIIT	ITKSLDREKQ	ARYEIVVEAR	DAQGLRGDSG	240
40	TATVLVTIQD	INDNFFPTFG	TKYTFVVPED	TRVGTSGVSL	FVEDPDEPQN	RMTKYSILRG	300
	DYQDAFTIET	NPAHNEGIIK	PMKPLDYEYI	QOYSFIVEAT	DPTIDLRYS	PPAGNRAQVI	360
	INITDWERPF	IFQPPFYHQ	LKENQKICPLI	GTVLAMPDPA	ARHSIGYSIR	RTSDKGQFFR	420
	VTKKGDIYNE	KELDREVYPW	YNLTVEAKEL	DSTGTPTGKE	SIVQVHIEVL	DENDNAPEFA	480
	KPYQPKVCEN	AVHQQLVLQI	SAIDKIDITPR	NVKFKFTLNT	ENNFTLTDNH	DNTANITVKY	540
45	GQFDREHTIG	HFLPVIISDN	GMPSTRGTST	LTVAVCKCNE	QGEPTFCEDM	AAQGVSVIQA	600
	VVAILLCLIT	ITVITLILFL	RRRLRKQARA	HGKSVPEIHE	QLVTYDEEGG	GEMDTTSYDV	660
	SVLNSVTRHG	AKPFRPALDA	RPSLYAQVQK	PPRHAPGAHG	GGGEMAMIE	VKKDEADHDG	720
	DGPPFYDTLHI	YGYGESIESA	ESLSSLTGDS	SDSDVDYDFL	NDWGPFRKML	AELYGSDPRE	780
50	ELLY						

Seq ID NO: 340 DNA sequence
Nucleic Acid Accession #: NM_003088
Coding sequence: 112-1593

	1	11	21	31	41	51	
55	GCGGAGGGTG	CGTGCGGGCC	GCGGCGAGCG	AACAAAGGAG	CAGGGGGGCC	GCCGCGAGGA	60
	CCGCGCACCC	ACCTCCCGGG	GCGGCGAGCG	GCGCTCTCGT	CTACTGCCAC	CATGACCGCC	120
60	AACGGCACAG	CCGAGGCGGT	GCAGATCCAG	TTCGGGCTCA	TCAACTCGCG	CAACAAGTAC	180
	CTGACGGCGG	AGGCGTTCGG	GTTCAAGGTG	AACGCGTCCG	CCAGCAGCCT	GAAGAAGAAG	240
	CAGATCTGGA	CGCTGGAGCA	GCCCGCTGAC	GAGGCGGGCA	GCGCGGCGGT	GTCCTGCGC	300
	AGCCACCTGG	GCCGCTACCT	GCGGCGGAGC	AAGGACGGCA	ACGTGACCTG	CGAGCGCGAG	360
	GTGCGCGGTC	CCGACTGCGC	TTTCTCTATC	GTGGCGCAOG	ACGACGCTCG	CTGGTCGCTG	420
65	CAGTCCGAGG	CGCACCGGCG	CTACTTCGGC	GGCACCGAGG	ACCGCTCTGC	CTGCTTCGCG	480
	CAGACGGTGT	CCCCCGCCGA	GAGTGGAGC	GTGCACATCG	CCATGCACCC	TCAGTCAAC	540
	ATCTACAGTG	TACCCGTA	GCGCTAGCGG	CACCTGAGCG	CAGGCGCGGC	CGACGAGATC	600
	GCGGTGGACC	GCGACGTGCC	CTGGGGGCTC	GACTCGCTCA	TCACCTCGC	CTTCCAGGAC	660
	CAGCGCTACA	GCGTGCAGAC	GCGCGACAC	CGCTCTCTGC	GCCACGACGG	GCGCTGGTGG	720
70	GCGGCGCCCG	AGCGGCGCAC	TGGCTACACG	CTGGAGTTCC	GCTCCGCGAA	GGTGGCCTTC	780
	CGCGACTGCG	AGGGCCGTTA	CCTGGCGCGG	TGCGGGCCCA	GCGGCGCGCT	CAAGCGGGCG	840
	AAGGCCACCA	AGGTGGGCAA	GGACGAGCTC	TTTGCTCTGG	AGCAGAGCTG	CGCCGAGGTC	900
	GTGCTGCAGG	CGGCCAACGA	GAGGAACGTG	TCCACGCGCC	AGGGTATGGA	CCTGTCTGCC	960
75	AATCAGGAGC	AGGAGACCGA	CCAGGAGACC	TTCCAGCTGG	AGATCGACCG	CGACACCAAA	1020
	AAGTGTGCTT	TCCGTACCCA	CACGGGCAAG	TACTGGACGC	TGACGGCCAC	CGGGGGCGTG	1080
	CAGTCCACCG	CCTCCAGCAA	GAATGCCAGC	TGCTACTTTG	ACATCGAGTG	GCGTGACCGG	1140
	CGCATCACAC	TGAGGGCGTC	CAATGGCAAG	TTTGTGACCT	CCAAGAAGAA	TGGGCGAGTG	1200
	GCGGCTCGGG	TGGAGACAGC	AGGGGACTCA	GAGCTCTTCC	TCATGAAGCT	CATCAACCGC	1260
	CCCATCATCG	TGTTTCGCGG	GGAGCATGGC	TTTCATCGGT	GCGGCAAGGT	CACGGGCAAC	1320
80	CTGGAGGCCA	ACCGCTCCAG	CTATGACGTC	TTCCAGCTGG	AGTTCAACGA	TGGGCGCTAC	1380
	AACATCAAAG	ACTCCACAGG	CAAATACTGG	ACGGTGGGCA	GTGACTCCGC	GGTCAACAGC	1440
	AGCGGGACCA	CTCCTGTGGA	CTTCTTCTTC	GAGTCTCTGG	ACTATAACAA	GGTGGCCATC	1500
	AAGGTGGGCG	GGGCTTACCT	GAAGGGGAGC	CACGCGAGCG	TCTGGAAGGC	CTCGGCGGAA	1560
	ACGCTGGACC	CGGCTCTGCT	CTGGGAGTAC	TAGGGCGGCG	CGGCTCTTCC	CGGCGGCTGC	1620
85	CCACATGGCG	GCTCCTGCCA	ACCTCTCTCG	CTAACCCCTT	CTCGGCGAGG	TGGGCTCCAG	1680
	GGCGGGAGGC	AAGCCCCCTT	GCCTTTCAAA	CTGGAACACC	CAGAGAAAC	GGTGCCCCCA	1740
	CCTGTGCGCC	CTATGGAGTC	CCCACTCTCC	CCTCGCGCGG	GGTTCCTTAC	TCCCTCTGGG	1800

5	TCAGCGGCTG	CGGCGTGGCC	CTGGGAGGGA	TTTCAGATGC	CCCTGCCCTC	TTGTCTGCCA	1860
	CGGGGCGAGT	CTGGCACCTC	TTTCTTCTGA	CCTCAGACGG	CTCTGAGCCT	TATTTCTCTG	1920
	GAAGCGGCTA	AGGAGCGGTT	GGGGGCTGGG	AGCCCTGGGG	GTGTAGTGTA	ACTGGAATCT	1980
	TTTGCTCTTC	CCAGCCACCT	CCTCCAGGCC	CCCCAGGAGA	GCTGGGCACA	TGTCCCAAGC	2040
	CTGTCACTGG	CCCTCCCTGG	TGCACTGTCC	CGAAACCCCC	TGCTTGGGAA	GGGAAGCTGT	2100
	CGGGAGGAGT	AGGACTGACC	CTTGTGGTGT	TTTTTTGGGT	GGTGGCTGGA	AACAGCCCTT	2160
	CTCCACAGTG	GGAGAGGCTC	AGCCTGGCTC	CCTTCCTTGG	AGCGGCAGGG	CGTGACGGCC	2220
	ACAGGGTCTG	CCCGCTGCAC	GTTCCTGCCA	GGTGGTGGTG	GCGGGCGGGT	AGGGGTGTGG	2280
10	GGGCGCTCTT	CCTCTGTCTT	CCTTCCTTTC	ACCCTAGCCT	GACTGGAAGC	AGAAAATGAC	2340
	CAAATCASTA	TTTTTTTTTAA	TGAAATATTA	TTGCTGGAGG	CGTCCAGGCC	AAGCCTGGCT	2400
	GTAGTAGGGA	GTGATCTGGC	GGGGGGCGTC	TCAGCACCTT	CCCCAGGGGG	TGCATCTCAG	2460
	CCCCCTCTTT	CCGTCTTTC	CGTCCAGCCC	CAGCCCTGGG	CCTGGGCTGC	CGACACCTGG	2520
	GCCAGAGCCC	TGCTGTGAT	TGGTGTCTCC	TGGGCTCTCC	GGGTGGATGA	AGCCAGGCGT	2580
15	CGCCCCCTCC	GGGAGCCCTG	GGGTGAGCG	CGGGGGCCCC	CCTGCTGCCA	GCCTCCCCCG	2640
	TCCCCAAT	GCATCTCACT	CTGGGTGTCT	TGGTCTTTTA	TTTTTTGTAA	GTGTCAATTG	2700
	TATAACTCTA	AACGCCCATG	ATAGTAGCTT	CAAACTGGAA	ATAGCGAAAT	AAAATAACTC	2760
	AGTCTGC						

20 Seq ID NO: 341 Protein sequence
Protein Accession #: NP_003079

25	1	11	21	31	41	51	
	MTANGTAEAV	QIQFGLINCG	NKYLTAEPFG	FKVNASASSL	KKKQIWTLEQ	PPDEAGSAAV	60
	CLRSHLGRYL	AADKDGNVTC	EREVPGPDCR	FLIVAHDDGR	WSLQSEAHRR	YFGGTEDLRS	120
	CPAQTVSPA	KWSVHIAHMP	QVNIYSVTRK	RYAHLRSARPA	DEIAVD RDVP	WGVDSLITLA	180
	FQDQRYSVQT	ADHRFLRHDG	RLVARPEPAT	GTYLEFRSGK	VAFRDCEGRY	LAPSGPSGTL	240
	KAGKATVVGK	DELFLALQSC	AQVVLQAANE	RNVSTRQMGD	LSANQDEETD	QBTFOLEIDR	300
	DTKKCAFRT	TGKYWTLTAT	GGVQSTASSK	NASCYFDIEW	RDRRITLRAS	NGKFVTSKKN	360
30	GQLAASVETA	GDSEFLFMKL	INRPIIVFRG	EHGFIGCRKV	TGTL DANRSS	YDVPQLEPND	420
	GAYNKDSTG	KWYTVGSDSA	VTSSGDTFVD	FFFEPCDYNK	VAIKVGGRYL	KGDHAGVLKA	480
	SAETVDPASL	WEY					

35 Seq ID NO: 342 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 660..1705

40	1	11	21	31	41	51	
	CGCTCCGCAC	ACATTTCTGT	TGCGGGCCTA	AGGGAAACTG	TTGGCGGCTG	GGCCCCCGGG	60
	GGGATTCTTG	GCAGTTGGGG	GGTCCGTCCG	GAGCGAGGGC	GGAGGGGAAG	GGAGGGGAA	120
	COGGTTGGG	GAAGCCAGCT	GTAGAGGGCG	GTGACCGCGC	TCCAGACACA	GCTCTGCTC	180
	CTCAGAGCGG	ACAGATCCAA	GTGGGAGCA	GCTCTGCGTG	CGGGGCTCA	GAGAATGAGG	240
45	COGGGTTTGG	CCCTGTGCTT	CCTCTGGCAG	CGCTCTGGC	COGGGCGGGG	CGCGCGGAA	300
	CACCCACTGA	CGGACCGTGC	TGGCTGCTCG	GCCTCGGGGG	CCTGTCTACG	CCTGCACCCAC	360
	GCTACCATGA	AGCGGCAGGC	GGCCGAGGAG	GCCTGCATCC	TGCGAGGTGG	GGCGCTCAGC	420
	ACCGTGGCTG	CGGGGCGCGA	GCTGGCGCGT	GTGCTCGCGC	TCTGTGCGGG	AGGCCCCAGG	480
	CCCGAGAGGG	GTCTCAAGA	CCTGTGTTTC	TGGGTGCGAC	TGGAGGCGAG	CGGTTCCAC	540
50	TGCACTCTG	AGAAAGAGCC	TTTGGGGGT	TTCTCTGGC	TGTCTCTCGA	CCCGGCGGT	600
	CTCGAAGCG	ACACGCTGCA	GTGGGTGGAG	GAGCCCCAAC	GCTCTGCTAC	CGCGGCGAGA	660
	TGCGCGTAC	TCCAGGCCAC	CGTGGGGTTC	GAGCCCGCAG	CTGGAAGGAG	ATGCGATGCC	720
	ACCTGGCGGC	CAACGGCTAC	CTGTGCAAGT	ACCACTTTGA	GGTCTTGTGT	CCTGTGCGCG	780
	GGCCCGGGGG	CGCCTCTAAC	TTGAGCTATC	GGCGCGCTTC	CCAGCTGCAC	AGCGCGGCTC	840
55	TGGACTTCAG	TCCACTTGGG	ACCGAGGTGA	GTGCGCTCTG	COGGGGACAG	CTCCGATCTC	900
	CAGTTACTTG	CATCGGCGAC	GAAATCGGCG	CTCGCTGGGA	CAAACTCTCG	GGCGATGTGT	960
	TGTGTCCCTG	CCCCGGGAGG	TACCTCCGTG	CTGSCAAATG	CGCAGAGCTC	CCTAACTGCC	1020
	TAGAGACTTT	GGGAGGCTTT	GCCTGCGAAT	GTGCTACGGG	CTTCGAGCTG	GGGAAGGACG	1080
60	CGCGCTCTTG	TGTGACCACT	GGGGAAGGAC	AGCCGACCTT	TGGGGGGACC	GGGGTGCCCA	1140
	CCAGGCGCCC	GGCGGCCACT	GCAACCAAGC	CGGTGCGCGA	GAGAACATGG	CCAATCAGGG	1200
	TGAGCGAGAA	GCTGGGAGAG	ACACCACTTG	TCCCTGAACA	AGACAATTCA	GTAACATCTA	1260
	TTCTGTAGAT	TCCTCGATGG	GGATCACAGA	GCAAGATGTC	TACCTTCAAA	ATGTCCCTTC	1320
	AAGCCGAGTC	AAAGGCCACT	ATCACCCCAT	CAGGGAGCGT	GATTTCCAAG	TTTAATTCTA	1380
	CGACTTCTTC	TGCCACTCCT	CAGGCTTTGG	ACTCCTCTCT	TGCCGTGGTC	TTCATATTGT	1440
65	TGAGCACAGC	AGTAGTAGTG	TTGGTGTATCT	TGACCATGAC	AGTACTGGGG	CTTGTCAAGC	1500
	TCTGCTTTCA	CGAAAGCCCC	TCTTCCACGC	CAAGGAAGGA	GTCTATGGGG	CGCGCGGGCC	1560
	TGGAGAGTGA	TCCTGAGCCC	GCTGCTTTGG	GCTCCAGTTC	TGCACATTGC	ACAAACAATG	1620
	GGGTGAAAGT	CGGGGACTGT	GATCTGCGGG	ACAGAGCAGA	GGGTGCTCTG	CTGGCGGAGT	1680
70	CCCTCTTGG	CTCTAGTGAT	GCATAG				

75 Seq ID NO: 343 Protein sequence
Protein Accession #: FGENESH predicted

75	1	11	21	31	41	51	
	MKGDFMTKTP	KAFATKAKID	KWDLIKLKSF	CTAKETIIRV	NSQPTDWQKT	FAIYPSDKGV	60
	IARIYKELEQ	IYKKKKPTKT	LRTHFLSRPK	GNCWPLGPRG	DSWQLGGPSG	ARABGKGGT	120
	GLGKPAVEGG	DRAPDTALRP	RAGQIQVGSS	SACGASENEA	GVRFPVPLAG	ALARAGRRT	180
	PHCRPCNLLG	LGGLLQAPAP	YHEAAGRGGG	LHPARWGAQH	RACGRRAARC	ARAPAGRPR	240
80	RRGLQRPAVL	GRTGQAQFPL	HPGERAFAGF	LLAVLRPRRS	RKRHAAVGGG	APTLLHRAEM	300
	RGTPGHRWGR	ARSWKEMRCH	LRANGYLCKY	QFEVLCPAPR	PGAASNLVSR	APPQLHSAAL	360
	DFSPPGTEVS	ALCRGLPLIS	VTCIADEIGA	RWDKLSGDVL	CPCPGRYLRA	GKCAELPNCL	420
	DDLGGFACBC	ATGPELGDGD	RSCVTSSEGG	PTLGGTGVP	RRPPATATSP	VPQRTPIRV	480
	DEKLGETPLV	PEQDMSVTSI	PEIPRWGSQS	TMSTLQMSLQ	ABSKATITPS	GSVISKFNST	540
85	TSSATPQAFD	SSSAVVFIFV	STAVVVLVIL	TMTVLGLVLK	CFHESPSSQG	RKESMGPPGL	600
	ESDPEPAALG	SSSAHCTNNG	VKVGDCDLRD	RABGALLAES	PLGSSDA		

Seq ID NO: 344 DNA sequence
Nucleic Acid Accession #: NM_012072
Coding sequence: 149-2107

5
0
5
20
25
30
35
40
45
50
55
60
65
70
75
80
85

1 11 21 31 41 51
AAAGCCTCA GCCTTTGTGT CCTTCTCTGC GCCGGAGTGG CTGCAGCTCA CCCCTCAGCT 60
CCCCTTGGGG CCCAGCTGGG AGCCGAGATA GAAGCTCCTG TCGCCGCTGG GCTTCTCGCC 120
TCCCGCAGAG GCCACACAGC AGACCGGATG GCCACCTCC ATGGGCGTGC TGCTGCTGCT 180
GCTGCTGCTC CTGACCCAGC CCGGGGCGGG GACGGGAGCT GACACGGAGG CGGTGGTCTG 240
CGTGGGGACC GCCTGCTACA CGGCCCACTC GGGCAAGCTG AGCGTGGCG AGGCCAGAA 300
CCACTGCAAC CAGAACGGGG GCAACCTGGC CACTGTGAAG AGCAAGGAGG AGGCCAGCA 360
CGTCCAGCGA GTACTGGCCC AGCTCCTGAG GCGGGAGGCA GCCCTGACGG CGAGGATGAG 420
CAAGTTCTGG ATTGGGCTCC AGCGAGAGAA GGGCAAGTGC CTGGACCTTA GTCTGCCGCT 480
GAAGGGCTTC AGCTGGGTGG GCGGGGGGGA GGACACGCTT TACTCTAACT GGCACAAGGA 540
GCTCCGGAAC TCGTGTACTT CCAAGCGCTG TGTGTCTCTG CTGCTGGACC TGTCCAGCC 600
GCTCCTTCCC AACCGCTGCG CCAAGTGGTC TGAGGGCCCC TGTGGGAGCC CAGGCTCCCC 660
CGGAAGTAAC ATTGAGGGCT TCGTGTGCAA GTTCAGCTTC AAAGGCATGT GCGGCTCT 720
GGCCCTGGGG GGGCCAGGTC AGGTGACCTA CACCACCCCC TTCCAGACCA CCAGTTCCTC 780
CTTGGAGGCT GTGCCCTTTG CCTCTGGGCG CAATGTAGCC TGTGGGGAAG GTGACAAGGA 840
CGAGACTCAG AGTCATTATT TCCTGTGCAA GGAGAAGGCC CCGCATGTGT TCGACTGGGG 900
CAGCTCGGGC CCCCTCTGTG TCAGCCCCAA GTATGGCTGC AACTTCAACA ATGGGGGCTG 960
CCACCGAGAC TGCTTTGAAG GGGGGGATGG CTCTTCTCTC TGCGGCTGCC GACCAGGATT 1020
CCGGCTGCTG GATGACCTGG TGACCTGTGC CTCTCGAAAC CCTTGCAGCT CCAGCCCATG 1080
TCGTGGGGGG GCCACGTGGC TCCTGGGACC CCATGGGAAA AACTACACGT GCCGCTGCCC 1140
CCAAGGTGAC CAGCTGGACT CGAGTCAGCT GGACTGTGTG GACCTGGATG AATGCCAGGA 1200
CTCCCCCTGT GCCCAGGAGT GTGTCAACAC CCTGGGGGCG TTCGCTGCG AATGTGGGT 1260
TGGCTATGAG CCGGGCGGTC CTGGAGAGGG GGCTGTCTAG GATGTGGATG AGTGTGCTCT 1320
GGGTGCGCTG CCTTGGCCCC AGGGCTGCAC CAACACAGAT GGCTCATTTT ACTGTCTCTG 1380
TGAGGAGGGG TACGTCTCTG CCGGGGAGGA CGGGACTCAG TGCCAGGAGG TGGATGAGTG 1440
TGTGGGCGCG GGGGGCCCCC TCTGGACAG CTGTGTCTTC AACACACAAG GGTCTTCCA 1500
CTGTGGCTGC CTGCCAGGCT GGGTCTGGC CCCAAATGGG GTCTCTTGA CAATGGGGCC 1560
TGTGTCTCTG GGACCCACAT CTGGGGCCCC CGATGAGGAG GACAAAGGAG AGAAAGAAAG 1620
GAGCACCGTG CCGCGGCTGC CAACAGCCAG TCCACAAGG GGGCCGAGG GCACCCCAA 1680
GGCTACACCC ACCACAAGT GACCTTCTGT GTCTCTGAC GCCCCATCA CATCTGCCCC 1740
ACTCAAGATG CTGGCCCCCA GTGGTCTCTC AGGCTCTGG AGGGAGCCCA GCATCCATCA 1800
CGCCACAGCT CCCTCTGGCC CCCAGGAGCC TGCAGGTGGG GACTCTCCG TGCCACACA 1860
AAACAAGCAT GGCACTGACG GGCAAAAGCT GCTTTTATTC TACATCTTAG GCACCGTGGT 1920
GGCCATCTTA CTCTGTCTGG CCCTGCTCTT GGGGCTACTG GTCTATCGCA AGCGGAGAGC 1980
GAAGAGGGAG GAGAAGAGG AGAAGAAGCC CCAGAATGCG GCAGACAGTT ACTCCTGGGT 2040
TCCAGAGCGA GCTGAGAGCA GGGCCATGGA GAACAGTAC AGTCCGACAC CTGGGACAGA 2100
CTGCTGAAAG TGAGGTGGCC CTAGAGACAC TAGAGTCACC AGCCACCATC CTCAGAGCTT 2160
TGAATCCCC ATTCCAAAGG GGCAACCCCA TTTTITGAA AGACTGGACT GGAATCTTAG 2220
CAACAATTTG TAAGTCTCCT CCTTAAAGGC CCCTTGGAAC ATGCAGGTAT TTCTACGGG 2280
TGTITTAGTG TCCTGAAGTG GAAGCTGTGT GTTGGCGTGC CACGGTGGGG ATTTCTGTAG 2340
TCTATAATGA TTGTACTTCC CCCTCCCTTT TCAAATCCA ATGTGACCAA TTCCGATCA 2400
GGGTGTAGG AGGCTGGGGC TAAGGGGCTC CCCTGAATAT CTCTCTGCT CACTTCCACC 2460
ATCTAAGAGG AAAAGGTGAG TTGCTCATGC TGATTAGGAT TGAATGATT TGTITCTCTT 2520
CTAGGATGA TAACTAAATC AATTAAATT TCAATTAGGT AAGAAGATCT GGTITTTTGG 2580
TCAAAGGAAA CATGTTCGGA CTGGAACAT TTCTTACAT TTGCATTCTT CCAATTCGCC 2640
AGCACAAAGT TTGCTAAATG TGATACTGTT GACATCTCC AGAATGGCCA GAAGTCAAT 2700
TAACTCTTA GGTGCAAGG AGGCAGGAAG TGCCCTTTTA GTTCTTACAT TTCTAATAGC 2760
CTTGGGTTTA TTGCAAAAGG AAGCTTGAAA AATATGAGAA AAGTTGCTTG AAGTGCATTA 2820
CAGGTGTTTG TAAGTCTACA TAATCTACGG GGCTAGGGCG AGAGAGGCCA GGGATTGTT 2880
CACAGATACT TGAATTAATT CATCCAAATG TACTGAGGTT ACCACACAT TGACTACGGA 2940
TGTGATCAAC ACTAACAAGG AAACAAATTC AAGGACAACC TGTCTTTGAG CCAGGGCAGG 3000
CCTCAGACAC CCGTCTGTG GCGCCGCTC CACTTCATCC TGCCCGGAAT GCCAGTCTC 3060
CGAGCTCAGA CAGAGGAAGC CCTGCAGAAA GTTCCATCAG GCTGTTTCTT AAAGGATG 3120
TGAACGGGAG ATGATGCACT GTGTTTGAAG AGTTGTCTAT TTAAGCATT TTAGCACAGT 3180
TCATAGTCCA CAGTTGATGC AGCATCCTGA GATTTTAAAT CCTGAAGTGT GGGTGGCGCA 3240
CACACCAAGT AGGGAGCTAG TCAGGCAAGT TGCTTAAAGG ACTTTTGTTC TCTGTCTCTT 3300
TTCTTAAAAA TTGGGGGTAA GGAGGGGAAG AAGAGGGAAA GAGATGACTA ACTAAATCA 3360
TTTTTACAGC AAAAAGTCTT CAAAGCCATT TAAATATAT CCTCATTTTA AAAGTTACAT 3420
TTGCAATAT TTCTCCCTAT GATAATGCAG TCGATAGTGT GCACTCTTTC TCTCTCTCTC 3480
TCTCTCTCAC ACACACACAC ACACACACAC ACACACACAC AGAGACACGG CACCATCTG 3540
CCTGGGGCAC TGGAACACAT TCCTGGGGGT CACCGATGGT CAGAGTCACT AGAAGTTACC 3600
TGAGTATCTC TGGGAGGCTT CATGTCTCCT GTGGGCTTTT TACCACCACT GTGCAGGAGA 3660
ACAGACAGAG GAAATGTGTC TCCTCCAAAG GCGCCAAAGC CTCAGAGAAA GGGTGTCTCT 3720
GGTTTTGGCT TAGCAATGCA TCGGTCTCTG AGGTGACACT CTGGAGTGGT TGAAGGGCCA 3780
CAAGGTGCAG GGTTAATACT CTGGCAGTT TTGAAATATA GATGCTATGG TTCAGATTGT 3840
TTTTAATAGA AAATAAAGG GGCAGGGGAA GTGAAAGGAA AGATGGAGGT TTTGTGCGGC 3900
TOGATGGGGC ATTTGGAATC TCTTTTAAAG GTCATCTCAT GGTCTCCAGT TTTCAAGTGG 3960
AACTCTGGTG TTTAACACTT AAGGGAGACA AAGGCTGTGT CCAATTGGCA AAACCTTCTT 4020
GGCCACGAGA CTCTAGGTGA TGTGTGAAGC TGGGCACTCT GTGGTGTGGA GAGCACCAT 4080
CTGTCTGGCC ATTCAAGAGA TTCTAAAGAC ATGGCTGGAT GCGCTGCTGA CCAACATCAG 4140
CACTTAAATA AATGCAAAATG CAACATTTCT CCCTCTGGGC CTTGAAAATC CTGTGCCCTT 4200
TCATTTGGGG TGAAGGAGAC ATTTCTGTCC TTGGCTTCCC ACAGCCCAA CAGGCTCTGT 4260
GTATGATCCC TGGGATCCAA CGAGCCCTCC TATTTTCACA GTGTTCTGAT TGCTCTCACA 4320
GCCCAGGCC ATGCTCTGTT CTCTGAATGC AGCCCTGTTC TCAACAACAG GGAGGTCAATG 4380
GAACCCCTCT GTGGAACCCA CAAGGGGAGA AATGGGTGAT AAAGAATCCA GTTCTCAAA 4440
ACCTTCCCTG CGAGGCTGGG TCCTCTCTCT GCTGGGTGGT GCTTTCTCTT GCACACCACT 4500
CCCAACAGG GGGGAGAGCC AGCAACCCAA CCAGACAGCT CAGGTGTGTC ATCTGATGGA 4560
AACCCTGGG CTCAACACG TGCTTTATTC TCCTGTTTAT TTTTGTCTGT ACTTTGAAGC 4620
ATGGAAATTC TTGTTGGGG GATCTTGGGG CTACAGTAGT GGGTAAACAA ATGCCCAACG 4680
GCCAAGAGGC CATTAAACAA TCGTCTTGT CCTGAGGGGC CCCAGCTTGC TCGGGCGTGG 4740
CACAGTGGG AATCCAAGGG TCACAGTATG GGGAGAGGTG CACCTGCCA CCTGCTAACT 4800

5
10
15
20
25
30
35

TCTCGTAGA CACAGTGTTC CTGCCCAGGT GACCTGTTC GCAGCAGAAC AAGCCAGGGC 4860
CATGGGAGC GGGGAAGTTT TCACCTGGAG ATGGACACCA AGACAATGAA GATTGTGTGT 4920
CCAAATAGGT CAATAATCTT GGGAGACTCT TGGAAAAAAC TGAATATATT CAGGACCAAC 4980
TCTCTCCCTC CCCTCATCCC ACATCTCAAA GCAGACAATG TAAAGAGAGA ACATCTCACA 5040
CACCCAGCTC GCCATGCCCTA CTCATTCCCTG AATTTCAGGT GCCATCACTG CTCCTTCTTT 5100
CTTCTTTGTC ATTTGAGAAA GGATGCAGGA GGACAATTCC CACAGATAAT CTGAGGAATG 5160
CAGAAAAAC AGGGCAGGAC AGTTATCGAC AATGCATTAG AACTTGGTGA GCATCCTCTG 5220
TAGAGGAGCT CCACCCCTGC TCAACAGCTT GGCTCCAGG CAAGACCAAC CACATCTGGT 5280
CTCTGCCCTC GGTGGCCAC ACACCTAAGC GTCATCGTCA TTGCCATAGC ATCATGATGC 5340
AACACATCTA CGTGTAGCAC TACGACGTTA TGTTTGGGTA ATGTGGGGAT GAACTGCATG 5400
AGGCTCTGAT TAAGGATGTG GGGAAAGTGGG CTGCGGTCAC TGTGGGCTT GCAAGGCCAC 5460
CTGGAGGCTT GTCTGTAGC CAGTGGTGGG GGAGCAAGGC TTCAGGAAGG GCCAGCCACA 5520
TGCCATCTTC CCTGCGATCA GGCAAAAAAG TGAATTAATA AAGTCAAAAC TTTATATGCA 5580
TGTGTTATGT CCATTTTGA GCATGAACG AGTTTAAAG AATTTTTTTT TCTCTCAAG 5640
TTGCTTTGTC TTTTCCATCC TCATCACAAG CCCCTGTGTT AGTGTCTTAT CCTGTAGCAA 5700
TCTTTCGATG GATGGAGATG ATCATTAGGT ACTTTGTGTT CAACCTTTAT TCTGTAAAT 5760
ATTTCTGTGA AACTAGGAG AACAGAGATG AGATTTGACA AAAAAAATT GAATTAATAA 5820
TAACACAGTC TTTTAAAC TAACATAGGA AAGCCTTTCC TATTATTCTT CTCTTAGCT 5880
TCTCCATTGT CTAAATCAGG AAAACAGGAA AACACAGCTT TCTAGCAGCT GCAAAATGGT 5940
TTAATGCCCC CTACATATTT CCATCACCTT GAACAATAGC TTTAGCTTGG GAATCTGAGA 6000
TATGATCCCA GAAACATCT GTCTCTACTT CGGCTGCAAA ACCCATGGTT TAAATCTATA 6060
TGGTTGTGTC ATTTTCTCAA CTAAAAATAG AGATGATAAT COGAATTCTC CATATATTCA 6120
CTAATCAAAG ACCTATTTTT CATACTAGAT TCCTGAGACA AATACTCACT GAAGGGCTTG 6180
TTTAAAAATA AATTGTGTTT TGGTCTGTTT TTGTAGATAA TGCCCTTCTA TTTTAGTAG 6240
AAGCTCTGTA ATCCCTTTAT TGTGCTGTTT CTCTTATCTG CAAGGTGGCA AGCAGTTCTT 6300
TTGAGCAGAT TTTGCCCATC ATTCCTCTGA GCTGAAGTTC TTGTCATAGA TTTGGCTTAA 6360
GCTTGAATTA GATCCCTGCA AAGGCTTGCT CTGTGATGTC AGATGTAATT GTAAATGTCA 6420
GTAATCACTT CATGAATGCT AAATGAGAAT GTAAGTATTT TTAATGTGT GTATTTCAAA 6480
TTTGTGTTGAC TAATCTGGA ATTACAAGAT TTCTATGCA GATTACCTT CATCCTGTGC 6540
ATGTTTCCCA AACTGTGAGG AGGGAAGGCT CAGAGATCGA GCTTCTCTC TGAGTTCTAA 6600
CAAAATGGTG CTTTGAGGGT CAGCCTTAG GAAGGTGCA CTTTGTGTC CTTTGAGCTT 6660
TCTGTTATGT GCCTATCCTA ATAACTCTT AAACACATT

Seq ID NO: 345 Protein sequence
Protein Accession #: NP_036204

40
45
50
55

1 11 21 31 41 51
| | | | |
MATSMGLLLL LLLLLTQPGA GTGADTEAVV CVGTACTYAH SGKLSAAEAQ NHCNQNGGNL 60
ATVKSKEEAQ HVQRVLAQLL RREALTARM SKFWIGLQRE KGKCLDPSLP LKGFVWVGGG 120
EDTPYENWHK ELRNSICISKR CVSLLLDLSQ PLLPNRLPKW SEGPCCGSPGS PGSNIEGFVC 180
KFSFGMCRPR LALGGPGQVT YTFPQTSS SLEAVPFASA ANVACGEGDK DETQSHYFLC 240
KEKAPDVFDW GSSGPLCVSP KYGCNPNNGG CHQDCPEBGD GSFLGCRPG FRLLDDLVTG 300
ASRNPCSSSP CRGGATCVLG PBKKNYTCRC PQGYQLDSSQ LDCVDVDEQ DSPCAQECVN 360
TPGGFRCECW VGYEPGGPGE GACQDVDECA LGRSPCAQGC TMTDGSFHCS CEEGVVLAGE 420
DGTQCQDVDE CVGPGGLCD SLCFNTQGSF HCGCLPGWVL APNGVSCVMG FVSLGPPSPG 480
PDEEDGKEKE GSTVPRATA SFTRGPBETP KATPTTSPRS LSSDAPITSA PLKMLAPSGS 540
SGVWREPSIH HATAASGPQE PAGGDSVSAT QNNDGTGQK LLLFYILGTV VAILLLLALA 600
LGLLVYRKRK AKREEKKKK PQNAADSYSW VPERAESRAM ENQYSPTPGT DC

Seq ID NO: 346 DNA sequence
Nucleic Acid Accession #: Z31560
Coding sequence: <1-966

55
60
65
70
75

1 11 21 31 41 51
| | | | |
CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCGGG CCGCAGCAA 60
ACTTCGGGGG GCGGGGGGG CAACCTCCACC GCGGGGGGG CCGGGGGCAA CCGAAAAAC 120
AGCCCGGACC GCGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGTTCCCG CCGGCGAGGG 180
CGCAAGATGG CCCAGAGAA CCCCAGATG CACAACCTCG AGATCAGCAA GCGCCTGGGG 240
GCCGAGTGA AACTTTTGTC GGAGACGAG AAGCGGCCGT TCATCGACGA GGCTAAGCGG 300
CTGCGAGCGC TGCACATGAA GGAGCACCG GATTATAAAT ACCGCGCCCG GCGGAAAAAC 360
AAGACGCTCA TGAAGAAGGA TAAGTACAG CTGCGCGGG GGCTGCTGGC CCCCGGCGGC 420
AATAGCATGG CGAGCGGGGT CGGGTGGGC GCGCGCTGG GCGCGGGCGT GAACACGCGC 480
ATGGACATG ACGCGCACAT GAACGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540
CAGCTGGGCT ACCGCGACGA CCGGGGCTC AATGCGCAG GCGCAGCGCA GATGCGAGCC 600
ATGCACGCT ACGACGTGAG CGCCCTGCAG TACAACCTCA TGACAGCTC CGAGACCTAC 660
ATGAACGGCT CGCCACCTA CAGCATGTCC TACTGCGAG AGGGCACCCC TGGCATGGCT 720
CTTGGCTCCA TGGGTTCCGT GGTCAAGTCC GAGGCGAGCT CCAGCCCCC TGTGTTACC 780
TCTTCTCTCC ACTCCAGGGC GCCCTGCCAG GCGGGGACC TCCGGGACAT GATCAGCATG 840
TATCTCCCCG GCGCGGAGGT GCGGGAACCC GCGGCCCCA GCAGACTTCA CATGTCCAG 900
CACTACCAGA GCGGCGCGGT GCCCGGCACG GCCATTAAAG GCACACTGCC CCTCTCACAC 960
ATGTGAGGGC CGGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAGAAAAA ACGAGGGAAA 1020
TGGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAAACCCG TACGCTCAAA 1080
AAAAA

Seq ID NO: 347 Protein sequence
Protein Accession #: CAA83435

80
85

1 11 21 31 41 51
| | | | |
HSARMYNNME TELKPPGPQQ TSGGGGNST AAAAGGNQKN SPDRVKRPMN AFMVWSRGQR 60
RKMAQENPKM HNSIISKRLG AEWKLLSETE KRPFIDEAKR LRALHMKRHP DYKYRPRRKT 120
KTLMKDKYT LPFGLLAPGG NSMASGVGVG AGLGAGVNR MDSYAHMNGW SNGSYSMQD 180
QLGYPQHPGL NAHGAQMOP MHRYDVSAQ YNSMTSSQTY MNGSPYSMS YSQQTGPGMA 240
LGSMSGVVKS EASSPPVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPVPEP AAPSRLLMSQ 300
HYQSGFVPGT AINGTLFLSH M

Seq ID NO: 348 DNA sequence
Nucleotide Accession #: NM_002638
Coding sequence: 120-473

5
1 11 21 31 41 51
CAATACAGCT AAGGAATTAT CCCTTGTAAT TACCACAGAC CGCCCTGGA GCCAGGCCAA 60
GCTGGAGCTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
10 TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGCTGGGACG CTGGTTCAG 180
AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240
TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300
CGCAAGAGCC AGTCAAAGAT CCAGTCTCCA CTAAGCCTGG CTCTGCCCC ATTATCTTGA 360
TCCGGTGGCG CATGTTGAAT CCCCTAACG GCTGCTTGA AGATACTGAC TGCCACAGAA 420
15 TCAAGAAGTG CTGTGAAGGC TCTTGGCGGA TGGCCTGTTT CGTCCCCAG TGAAGGGAGC 480
CGGTCTCTGC TGCACTCTGG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540
TGCTGCCCTT CCCCTTCCA CACTGTCCAT TCTTCTCTCC ATTCAAGATG CCCACGGCTG 600
GAGCTGCCTC TCTCATCCAC TTTCATAAA A

20 Seq ID NO: 349 Protein sequence:
Protein Accession #: NP_002629

25 1 11 21 31 41 51
MRASSFLIVV VFLIAGTLVL EAAVTGVFVK QQDVTVKGRVP FNGQDPVKQ VSVKQGDQVK 60
AQEPVKGPVS TKPGRSCPIIL IRCAMLNPPN RCLKDITDCPG IKKCEGSCG MACFVPQ

30 Seq ID NO: 350 DNA sequence
Nucleic Acid Accession #: NM_007183
Coding sequence: 75-2468

35 1 11 21 31 41 51
GAATTCGGA CAGGACGTGA AGATAGTTGG GTTTGGAGGC GGCCGCCAGG CCCAGGCCCG 60
GTGGACCTGC CGCATGTCAG GACGGTAACT TCCTGCTGTC GGCCCTGCAG CCTGAGGCCG 120
GCGTGTGCTC CCTGGGCTGG CCTCTGACC TGCACTGGA CCGCCGGGGC GCCGAGGGGC 180
CGGAGGCCGA GCGCTGCGG GCAGCCCGCG TCCAGGAGCA GGTCCGCGCC CGCCTCTTGC 240
40 AGCTGGGACA GCAGCCGGGG CACAACGGGG CCGCTGAGCC CGAGCCTGAG GCCGAGACTG 300
CCAGAGGCAC ATCCAGGGGG CAGTACCACA CCTGCGAGG TGGCTTCAGC TCTCGCTCTC 360
AGGCGCTGAG TGGGACAAG ACCTCGGGCT TCCGCCCCAT CGCCAAGCCG GCCTACAGCC 420
CAGCCTCTG GTCTCTCCCG TCCGCGGTGG ATCTGAGCTG CAGTGGAGG CTGAGTTTCA 480
CCCACAATGG GGGCAGCGCC TTTGGGGCCG CTGGGTACGG GGTGCCCCAG CCCACCCCTC 540
CCATGCCCCA CAGCCCGCTG TCCTTCCATG AGCGCGGTGG GGTGGGAGC CGGGCCGACT 600
45 ATGACACACT CTCTGTGCGG TCGCTGCGGG TGGGGCCCGG GGGCCTGGAC GACCGCTACA 660
GCTGTGTGTC TGAGCAGCTG GAGCCCGCGG CCACTCCAC CTACAGGGCC TTTGCGTACG 720
AGCGCCAGGC CAGCTCCAGC TCCAGCCGGG CAGGGGGGCT GGACTGGCCC GAGGCCACTG 780
AGGTTTCCCC GAGCCCGACC ATCCGTGCCC CTGCGGTGGG GACCTGTCAG CGATTCCAGA 840
GCAGCCACCG GAGCCCGGGG GTAGGCGGGG CAGTGCCCGG GGCCGTCTG GAGCCAGTGG 900
50 CTCGAGCGCC ATCTGTGCGG AGCCTCAGCC TCAGCCTGGC TGACTCGGGC CACTGCGCGG 960
ACGTGATGAG GTTCAACAGC TACGGTAGCC ACCGAACCTT GCAGAGACTC AGCAGCGGTT 1020
TTGATGACAT TGACTTGCCC TCAGCAGTCA AGTACCTCAT GGCTTCAGAC CCCAACCTGC 1080
AGGTGCTGGG AGCGGCTTAC ATCCAGCACA AGTGTCTACG CGATGCAGCC GCCAAGAAGC 1140
AGGCCCGCAG CCTTCAGGCC GTGCTTAGGC TGGTGAAGCT CTTCAACCAC GCCAACCAGG 1200
55 AAGTGCAGCG CCAATGCCCA GGTGCCATGC GCAACCTCAT CTACGACAAC GCTGACAACA 1260
AGCTGGCCCT GGTGAGGAG AAOGGGATCT TCGAGCTGCT GGGGACACTG GGGGAGCAGG 1320
ATGATGAGCT TCGCAAAAT GTCAAGGGA TCCTGTGGA CCTTTCATCC AGCGACCACC 1380
TGAAGGACCG CCTGGCCAGA GACACGCTGG AGCAGCTCAC GGACCTGGTG TTAGGCCCCC 1440
TGTGCGGGGC TGGGGGTCCC CCCCTCATCC AGCAGAACCG CTCGGAGGCG GAGATCTTCT 1500
60 ACAACGCCAC CGGCTTCTCT AGGAACCTCA GCTCAGCCTC TCAGGCCACT GCCCAGAAGA 1560
TGCGGGAGTG CCACGGCTG GTGGAAGCCC TGGTCACTTC TATCAACCAC GCCCTGGACG 1620
CGGGCAATG CGAGGACAAG AGCGTGGAGA ACGCGGTGTG CTCTCTGCGG AACCTGTCCT 1680
ACCGCTCTA CGACGAGATG CCGCGTCCG CGCTGCAGCG GCTGGAGGGT CGCGGCCGCA 1740
GGGACCTGCG GGGGGCGCGG CCGGGAGAGG TCGTGGGCTG CTTACGCGCG CAGAGCCGCG 1800
65 GGTGCGCGA GCTGCCCTC GCCGCGATG CGCTCACCTT CGCGGAGGTG TCCAAGGACC 1860
CCAAGGGCTC CAGTGGCTG TGGAGCCCC AGATCGTGGG GCTGTACAAC CGGTGCTGCT 1920
AGCGCTGCGA GCTCAACCG CACACGACCG AGCGGCGCG CGGGCGCTG CAGAACATCA 1980
CGGCAGGCGA CGCAGGTGG GCGGGGGTGC TGAGCCGCTT GGCCCTGGAG CAGGAGCGTA 2040
TTCTGAACCC CCTGTAGAC CGTGTAGGA CCGCCGACCA CCACAGCTG CGCTCACTGA 2100
70 CTGGCTCAT CCGAAACCTG TCTCGGAAAG CTAGGAACAA GGACGAGATG TCCACGAAGG 2160
TGGTGAAGCA CCTGATCAG AAGCTGCCAG GCAGCGTGGG TGAGAAGTGG CCCCAGCCG 2220
AGGTGCTGGT CAACATCATA GCTGTGCTCA ACAACCTGGT GGTGGCCAGC CCCATCGCTG 2280
CCGAGACCT GCTGTATTTT GACGGACTCC GAAAGCTCAT CTTTCATCAG AAGAAGCGGG 2340
75 ACAGCCCGGA CAGTGAAGG TCCTCCCGGG CAGCATCCAG CCTCTGCCC AACCTGTGGC 2400
AGTACAACAA GCTCCACCGT GACTTTCGGG CGAAGGGCTA TCGGAAGGAG GACTTCTTGG 2460
GCCCATAGGT GAAGCCTTCT GGAGGAGAAG GTGAGTGGC CCAGCGTCCA AGGGACAGAC 2520
TCAGCTCCAG GCTGCTTGGC AGCCACGCTT GGAGGAGAAG GCTAATGACG GAGGGGCCCC 2580
TCGCTGGGGC CCTGTGTGTC ATCTTTGAGG GTCTGGGGCC ACCAGGAGGG CGAGGGTCTT 2640
80 ATAGCTGGGG ACTTGTCTTC CGCAGGGCAG GGGGTGGGGC AGGGCTCAAG GCTGTCTTGG 2700
TGATAGGGGT GGTGACCCAG TCACATTGGC AGAGGTGGGG GTTGGCTGTG GCCTGGCAST 2760
ATCTTGGGAT AGCCAGCACT GGGATAAAG ATGGCCATGA ACAGTCAACA AAAAAAAAAA 2820
AAAAGGAATT C

85 Seq ID NO: 351 Protein sequence:
Protein Accession #: NP_009114.1

1 11 21 31 41 51

5
10
15

MQDGNFLLSA	LQPEAGVCSL	ALPSDLQLDR	RGAEGPFAER	LRAARVQEQV	RARLLQLGQQ	60
PRHNGAAEPE	PEAETARGTS	RQGYHTLQAG	FSSRSQGLSG	DKTSGFRPIA	KPAYSPASWS	120
SRSAVDLSCS	RLSSAHNGG	SAPGAAGYGG	AQPTPFMPTR	PVSFHERGGV	GSRADYDTLS	180
LSRLRLPGG	LDRLSLVSE	QLEPAATSTY	RAFAYERQAS	SSSRAGGLD	WPEATEVSPS	240
RTIRAPAVRT	LQRFQSSHRS	RGVGGAVPGA	VLEPVARAPS	VRSLSLSLAD	SGHLPDVHGF	300
NSYGSHTLQ	RLSSGFDDID	LPSAVKYLMA	SDPNLQVLGA	AYIQHKCYSD	AAAKQARSLL	360
QAVPRLVKLF	NHANQEVQRH	ATGAMRNLIY	DNADNKLALV	EENGIFELLR	TLREQDDELRL	420
KNVTGILWNL	SSSDHLKDRL	ARDTLEQLTD	LVLSPLSGAG	GPPLIQONAS	EABIFYNATG	480
FLRNLSSASQ	ATRQKMRCH	GLVDALVTSI	NHALDAGKCE	DKSVENAVCV	LRNLRYRLVD	540
EMPPSALQRL	EGRGRRDLAG	APPGEVVGC	TPQSRRLREL	PLAADALTFA	EVSKDPKGLG	600
WLWSFQIVGL	YNRLQLRCEL	NRHTTEAAAG	ALQNTAGDR	RWAGVLSRLA	LEQERILNPL	660
LDVRVTADHH	QLRSLTGLIR	NLSRNRNARD	EMSTKVVSHL	IEKLPGSVGE	KSPPAEVLVN	720
IIAVLNLLV	ASPIAARDLL	YFDGLRLKLF	IKKKRDSPTS	EKSSRAASSL	LANLWQYNKL	780
HRDFRAKGYR	KEDFLGP					

Seq ID NO: 352 DNA sequence
Nucleic Acid Accession #: M31469
Coding sequence: 1-651

20
25
30
35

1	11	21	31	41	51	
ATGGCTCGCC	AGGGAGAGCC	CCAGGTCCAG	TTCAAACCTG	TATTGGTTGG	TGATGGTGGT	60
ACTGGAAAAA	CGACCTTCGT	GAAACGTCT	TTGACTGGTG	AATTTGAGAA	GAAGTATGTA	120
GCCACCTTGG	GTGTTGAGGT	TCATCCCTTA	GTGTTCCACA	CCAACAGAGG	ACCTATTAAAG	180
TTCAATGTAT	GGGACACAGC	CGGCCAGGAG	AAATTCGGTG	GACTGAGAGA	TGCTATTAT	240
ATCCAAGCCC	AGTGTGCCAT	CATAATGTTT	GATGTAACAT	CGAGAGTTAC	TTACAAGAAT	300
GTGCCTAACT	GGCATAGAGA	TCTGGTACGA	GTGTGTGAAA	ACATCCCCAT	TGTCTTGTGT	360
GGCAACAAAG	TGGATATTAA	GGACAGGAAA	GTGAAGGCGA	AATCCATTGT	CTTCCACCGA	420
AAGAAGAATC	TTCACTACTA	CGACATTCTT	GCCAAAAGTA	ACTACAACCT	TGAAAAGCCC	480
TTCTCTGCGC	TTGCTAGGAA	GCTCATTGGA	GACCCTAAC	TGGAATTTGT	TGCCATGCCT	540
GCTCTGCCCC	CACCAAGAGT	TGTCATGGAC	CCAGCTTTGG	CAGCACAGTA	TGAGCACGAC	600
TTAGAGGTTG	CTCAGACAAC	TGCTCTCCCG	GATGAGGATG	ATGACCTGTG	A	

Seq ID NO: 353 Protein sequence
Protein Accession #: AAA36546

40
45

1	11	21	31	41	51	
MAAQGEPOVQ	FKLVLVGDDG	TGKTFVKKR	LTGEFEKKYV	ATLGVEVHPL	VFHTNRGPIK	60
FNVDWTAGQE	KFGGLRDGY	IQACAIIMP	DVTSRVTYKN	VPNWRDLVR	VCEINPIVLC	120
GNKVDIKDRK	VKAKSIVFHR	KKNLQYYDIS	AKSNYNFEKP	PLWLARKLIG	DPNLEFVAMP	180
ALAPEVVMVD	PALAAQYEH	LEVAQTALP	DEDDDL			

Seq ID NO: 354 DNA sequence
Nucleic Acid Accession #: NM_002820
Coding sequence: 304-831

50
55
60
65
70
75
80

1	11	21	31	41	51	
CGGGTTGCGA	AAGAAGCTGA	CTTCAGAGGG	GGAAACTTTC	TTCTTTTAGG	AGGCGGTTAG	60
CCCTGTTCCA	CGAACCCAGG	AGAACTGCTG	GCCAGATTAA	TTAGACATTG	CTATGGGAGA	120
CGTGTAAACA	CACCTACTAT	CATTGATGCA	TATATAAAAC	CATTTTATTT	TOGCTATTAT	180
TTCAAGAGAA	GGCCTCTGTA	TTTGTTCTTT	TTTTCCCTTT	TTGCTCTTTC	TGGCTGTGTG	240
GTTTGGAGAA	AGCACAGTTG	GAGTAGCCGG	TTGCTAAATA	AGTCCCGAGC	GCGAGCGGAG	300
ACGATGCAGC	GGGACTGGT	TCAGCAGTGG	AGCGTGGCGG	TGTTCTGCTG	GAGCTACGCG	360
GTGCCCTCCT	GGCGGCGCTC	GGTGGAGGTT	CTCAGCCGCC	GCCTCAAAAG	AGCTGTGTCT	420
GAACATCAGC	TCTCTCATGA	CAAGGGGAAG	TCCATCCAAG	ATTTACGGCG	ACGATTCTTC	480
CTTCAACATC	TGATGCGAGA	AATCCACACA	GCTGAAATCA	GAGCTACCTC	GGAGGTGTCC	540
CCTAACTCCA	AGCCCTCTCC	CAACACAAAG	AACCACCCCG	TCCGATTTGG	GTCTGATGAT	600
GAGGGCAGAT	ACCTAACTCA	GGAACTAAC	AAGGTGGAGA	CGTACAAAGA	GCAGCCGCTC	660
AAGACACCTG	GGAAAGAAAA	GAAAGGCAAG	CCCGGGAAAC	GCAAGGAGCA	GGAAAAGAAA	720
AAACGGCGAA	CTCGCTCTGC	CTGGTTAGAC	TCTGGAGTGA	CTGGGAGTGG	GCTAGAAGGG	780
GACCACTCTG	CTGACACCTC	CACAACGTCG	CTGGAGCTCG	ATTACAGGTA	ACAGGCTTCT	840
CTGGCCCGTA	GCCTCAGCGG	GGTGCTCTCA	GCTGGGTTTT	GGAGCCTCCC	TTCTGCCTTG	900
GCTTGGACAA	ACCTAGAAAT	TTCTCCCTTT	ATGTATCTCT	ATCGATTGTG	TAGCAATTGA	960
CAGAGAATAA	CTCAGAAAT	TGCTGCTCT	AAAGCAGTAC	CCCCCTACCA	CACACACCCC	1020
TGTCTCCAG	CACCATAGAG	AGGCGCTAGA	GCCCATTCCT	CTTCTCCAC	CGTCACCCAA	1080
CATCAATCCT	TTACCACTCT	ACCAATAAT	TTCATATTCA	AGCTTCAGAA	GCTAGTGACC	1140
ATCTTCATAA	TTTGCTGGAG	AAGTGTATTT	CTTCCCTTCA	CTCTCACACC	TGGGCAAACT	1200
TTCTTCAGTG	TTTTCATT	CTTACGTTCT	TTCACTTCAA	GGGAGAAATAT	AGAAGCAATT	1260
GATATTATCT	ACAAACACTG	CAGAACAGCA	TCATGTCATA	AACGATTCTG	AGCCATTAC	1320
ACTTTTATT	TAATTAATG	TATTTAAATTA	AATCTCAAAT	TTATTTTAAT	GTAAAGAACT	1380
TAAATTATGT	TTTAAACACA	TGCCTTAAAT	TTGTTTAAAT	AAATTTAACT	CTGGTTTCTA	1440
CCAGCTCATA	CAAAATAAAT	GGTTTCTGAA	AATGTTTAA	TATTAACCTA	CAAGGATATA	1500
GGTTTCTTCT	ATGTATCTTT	TGTTTCTATG	GCAAGATGAA	ATAATTTTTC	TAGGGTAATG	1560
CCGTAGGAAA	AATAAACTT	CACATTTAAA	AAAAA			

Seq ID NO: 355 Protein sequence
Protein Accession #: NM_002820

85

1	11	21	31	41	51	
MQRRLVQQWS	VAVFLLSYAV	PSGGRSVEGL	SRRLKRAVSE	HQLLHDKGKS	IQDLRRRFFL	60
HHLIAEIHIA	EIRATSEVSP	NSKPSPNTRN	HPVRFGSDDE	GRYLTQETNK	VETYKEQPLK	120

TPGKKKKGKP GRRKEQEKKK RRTRSWLDS GVTGSGLEGD HLSDTSTTSL ELDSR

Seq ID NO: 356 DNA sequence
Nucleic Acid Accession #: NM_017522
Coding sequence: 1-2100

5

1 11 21 31 41 51
10 ATGGGCTTCC COGAGCCGGG CCTCTCCGGG CTTCTGGGCG TGCTGCTGCT GCTGCTGCTG 60
CTGCTGCTGCT TCGGCTCCCA GCATCTTGGG GCGGCGAGCG CTGATCCGCT GCTCGGCGGC 120
CAAGGGCCGG CCAAGGAGTG CGAAAAGGAC CAATTCCAGT GCCGGAACGA GCGCTGCATC 180
CCCTCTGTGT AGAGATGGGA CGAGGACGAT GACTGCTTAG ACCACAGCGA CGAGGACGAC 240
TGCCCAAGA AGACCTGTGC AGACAGTGAC TTCACTGTG ACAACGGCCA CTGCATCCAC 300
GAACGGTGGA AGTGTGACGG CGAGGAGGAG TGTCTGATG GCTCGATGA GTCGAGGCC 360
15 ACTTGACCCA AGCAGGTGTG TCCTGCAGAG AAGCTGAGCT GTGAGCCAC CAGCCACAAG 420
TGTTGACTCT CCTCGTGGCG CTGCGACGGG GAGAAGGACT GCGAGGGTGG AGCGGATGAG 480
GCCGGCTGTG CTACCTCACT GGGCACCTGC CGTGGGACG AGTTCAGTG TGGGGATGGG 540
ACATGTGTCC TTGCAATCAA GCATGCAAC CAGGAGCAGG ACTGTCCAGA TGGGAGTGAT 600
GAAGCTGGCT GCCTACAGG GCTGAACGAG TGTCTGCACA ACAATGGCGG CTGCTCACAC 660
20 ATCTGCACTG ACCTCAAGAT TGGCTTTGAA TGCACTGCC CAGCAGGCTT CCAGCTCCTG 720
GACCAAGA CTGTGGCGA CATTGATGAG TGCAAGGACC CAGATGCCCTG CAGCCAGATC 780
TGTTGCAATG ACAAGGTGCT TTTAAGTGT GAGTGTCTACC CTGGCTGCGA GATGGACCTA 840
CTGACCAAGA ACTGCAAGGC TGCTGTGCGC AAGAGCCCAT CCTTAATCTT CACCAACCGC 900
ACGAGTGGCG AGGATCGACC TGTGAAGCGG AACTATTAC CCTCATCCC CATGCTCAAG 960
25 AATGTCTGTG GAATGAGTG GGAAGTTGCC ACCAATCGCA TCTACTGGTG TGACCTCTCC 1020
TACCGTAAGA TCTATAGCGC CTACATGGAC AAGGCCAGTG ACCCGAAGA GCGGGAGGTC 1080
CTCATTGAGC AGCAGTTTCA CTCTCCAGAG GGCCTGGCAG TGGACTGGGT CCACAAGCAC 1140
ATCTACTGGA CTGACTCGGG CAATAGACC ATCTCAGTGG CCACAGTTGA TGGTGGCCGC 1200
CGACGCACTC TCTTCAGCGG TAACCTCAGT GAACCCCGGG CCATCGCTGT TGACCCCTCTG 1260
30 CGAGGGTTCA TGTATTGGTC TGACTGGGGG GACCAAGGCA AGATTGAGAA ATCTGGGCTC 1320
AAGCGTGTGG ACCGGCAAC ACTGGTGTCA GACAATATTG AATGGCCCAA CGGAATCACC 1380
CTGGATCTGC TGAGCCAGCG CTTGTACTGG GTAGACTCCA AGCTACCACT ACTGTCCAGC 1440
ATTGACTTCA GTGAGGCGAA CAGAAAGACG CTGATCTCCT CCACTGACTT CTTGAGCCAC 1500
CCTTTTGGGA TAGCTGTGTT TGAGGACAAG GTGTTCTGGA CAGACCTGGA GAACGAGGCC 1560
35 ATTTTCAGTG CAATTCGGCT CAATGGCCTG GAAATCTCCA TCCTGGCTGA GAACCTCAAC 1620
AACCACATG ACATGTCTAT CTTCATGAG CTGAAGCAGC CAAGAGCTCC AGATGCTCTG 1680
GAGCTGAGTG TCCAGCTTAA TGGAGGCTGT GAATACCTGT GCCTCTCTGC TCCTCAGATC 1740
TCCAGCCACT CTCCCAAGTA CACATGTGCC TGTCTGACA CAATGTGGCT GGGTCCAGAC 1800
ATGAAGAGGT GGTACCGAGA TGCAATGAA GACAGTAAGA TGGGCTCAAC AGTCACTGCC 1860
40 GCTGTATCGT GATATCGTCT GCCATAGTGT GTGATAGCCC TCCTGTGCAT GAGTGATAC 1920
CTGATCTGGA GAAATCGGAA CGGGAAGAAC ACCAAAAGCA TGAATTTTGA CAACCCAGTC 1980
TACAGGAAAT CAACAGAAGA AGAAGATGAA GATGAGCTCC ATATAGGGAG AACTGCTCAG 2040
ATTGGCCATG TCTATCTGTC ACGAGTGGCA TTAAGCCTTG AAGATGATGG ACTACCTGTA 2100
GGATGGGATC ACCCCCTTGG TGCCCTATGG AATTCACTCC CATGCACTAC ACTCCGGATG 2160
45 GTGATGACT GGATGAATGG GTTTCTATAT ATGGGTCTGT GTGAGTGTAT GTGTGTGTGT 2220
GATTTTTTGT TTTAAATTTA TGTTCGGGAA AGGTAACCAC AAGTTATGA TGAATGCAA 2280
ACATCCAAAG GATGTGAGAG TTTTCTATG TATAATGTTT TATACACTTT TTAAGTGGTT 2340
GCACTACCCA TGAGGAATTC GTGGAATGGC TACTGCTGAC TAACATGATG CACATAACCA 2400
AATGGGGGCC AATGGCAGAG TACCTTACTC ATCATTATAA AACTATAATT ACAGAAGATG 2460
50 TTTGGTGTCT GGGGGGCTTT TTTAGGTTTT GGGCATTGT TTTTGTAAA TAAGATGATT 2520
ATGCTTTGTG GCTATCCATC AACATAAGT

Seq ID NO: 357 Protein sequence
Protein Accession #: NP_059992

55

1 11 21 31 41 51
55 MGLPEPGLR LLALLLLLLL LLLRLQLHL AAAADPLLGG QGPAKECEKD QFQCRNERCI 60
PSVWRCEDD DCILDHSEDD CPKKTCAAD FTCDNGHCIE ERWKCDGEEB CPDGSDESEA 120
60 TCTKQVCPAE KLSGCTPSHK CVPASWRCDG EKDCGEGADE AGCATSLGTC RGDEFQCGDG 180
TCVLAIKHCN QBDQCPDQSD EAGCLQGLNE CLHNGGCSH ICTDLKIGFE CTCFAGFQLL 240
DQKTCGDIDE CKDPDACSQI CVNYKGYFKC ECPYGCMDL LTKNCKAAAG KSPSLIFTNR 300
70 TSAEDRPVKR NYSRLIPMLK NVVALDVEVA TNRIYWCDS YRKIYSAYMD KASDPKEREV 360
LIDELQLHPE GLAVDWVHKH IYNTDSGNTK ISVATVDGGR RRTLFNRNLS EPRAIADPL 420
65 RGFMYSWDG DQAKIEKSLG NGVDRQTLVS DNIEWPNGIT LDLLSRLYLW VDSKLHLSS 480
IDFSGGNRKT LISSDFLSH PFGIAVFEDK VFWTDLENEA IFSANRLNL EISILAENLN 540
NPHDIVIFHE LKQPRAPDAC ELSVQPNNGC EYLCLPAPQI SSHSPKYTCA CPDTMWLSPD 600
MKRCYRDANE DSKMGSTVTA AVIGIIVPIV VIALLCMSGY LIWRNWKRN TKSMNFDNPV 660
YRKTEEREDE DELHIGRTAQ IGHVYPARVA LSEDDGLP

Seq ID NO: 358 DNA sequence
Nucleic Acid Accession #: M27826
Coding sequence: <1-503

75

1 11 21 31 41 51
75 AGCCCAAGAA ACATCTCACC AATTTCAAAT CTGATCTATT CGGCTTAGCG ACTGAAGATT 60
GACGCTGCCC GATCGCCTCG GAAGTCCCCC GGACCATCAC AGAAGCCGAG CTTGGGTAA 120
80 CTCTCAGAGT GGAGGGTAAG TCCATCCCCC GTTTAATCGA TACGGGGGCT ACCCACTCCA 180
CGTTGCCCTC TTTTCAAGGG CCGTTTCCC TTGCCCCCAT AACTGTTGTG GGTATTGAGC 240
GCCAAGCTTC AAAACCCCTG AAAACTCCCC CACTCTGGTG CCAACTTGA CAACACTCTT 300
TTATGCACTC TTTTGTAGT ATCCCCACCT GCCCACTTCC CTTATTAGGC CGAAATATT 360
TAACCAAAAT ATCTGCTTCC CTGACTATTG CTGGAGTACA GCTACATCTC ATTGCTGCC 420
85 TTCTTCCCAA TCCAAGCCT CTTTGTGTG CTCTAACATC CCCCAATAT CAGCCCTTAC 480
CACAAAGACT CCTTCACTG TAATCTCTCC CACTCTAGGT TCCACGCGG CCCTAATCC 540
CACTTGAAGC AGCCCTGAGA AACATGCCCC ATTTCTCTCT CATACACCC CCCAAAATT 600
TTGCGCGCTC CAACACTTCA ACACATTTT GTTTTATTG TCTTATTAA ATCAAGAAGC 660

AGGAATGTCA GGCCTCTGAG CCCAGGCCAG GCCATGCGAT CCCTGTGAC TTGCAAGTAT 720
 ACATCCAGAT GGCCTGAAGT AACTGAAGAT CCACAAAAGA AGTAAAAACA GCCTTAACTG 780
 ATGACATGCC ACCATTGTGA TTGTCTCTG CCCCAACCTA ACTGATCAAT GTACTTTTGA 840
 ATCTCCCCCA CCTTAAAGAA GGTTCCTTGT AATTCTCCCC ACCCTTGAGA ATGTACTTTG 900
 TGAGATCCAC CCTGCCCCAC CAGAGAAACA CCCCTTTGA TTGTAATTTT TTATTACCTT 960
 CCCAAATCCT ATAAACACAG CCCACCCCTA TCTTCTTCA CTGACTCTCT TTTCCGACTC 1020
 AGCCACCGGC ACCCAGGTGA AATAAACAGC TTTATTGCTC AC

Seq ID NO: 359 Protein sequence
 Protein Accession #: AAA65999

1 11 21 31 41 51
 PKKHLTNFKS DLFGLATEDW RCPPIASEVPW TITEAELRVT LTVEGKSIPC LIDTGATHST 60
 LPSFQGPVSL APITVVGIDG QASKPLKTPF LWCQLQHSF MHSFLVIPTC PLPLLGRNIL 120
 TKLSASLTIP GVQLHLIAAL LPNPKPPLCP LTSFQYQPLP QDLPSA

Seq ID NO: 360 DNA sequence
 Nucleic Acid Accession #: NM_001854
 Coding sequence: 162-5582

1 11 21 31 41 51
 AACCATCAAA TTAGAAGAA AAAGCCCTTT GACTTTTTC CCCTCTCCTT CCCCAATGGC 60
 TGTGTAGCAA ACATCCCTGG CGATACCTTG GAAAGGACGA AGTTGGTCTG CAGTCCGAAT 120
 TTGTGGGTTT GAGTTTCACAG TTGTGAGTGC GGGGCTCGGA GATGGAGCCG TGGTCTCTTA 180
 GGTGGAAAAC GAAACGGTGG CTCTGGGATT TCACCGTAAC AACCTCCGCA TTGACCTTCC 240
 TCTTCCAGC TAGAGAGGCT AGAGGAGCTG CTCCAGTTGA TGTAATAAAA GCACTAGATT 300
 TTCACAATT TCACAGAGGA ATATCAAAAA CAACGGGATT TTGCACAAAC AGAAAGNATT 360
 CTAAGAGCTC AGATACTGCT TACAGAGTTT CAAAGCAAGC ACAACTCAGT GCCCCAACAA 420
 AACAGTTATT TCCAGGTGGA ACTTTCCCAAG AAGACTTTTC AATACTATT ACAGTAAAAAC 480
 CAAAAAAGAG AATTCACTGT TTCTTTTAT CTATATATAA TGAGCATGGT ATTCAGCAAA 540
 TTGTGTGTGA GGTGGGAGA TCACCTGTTT TTCTGTTTGA AGACCACACT GGAACACCTG 600
 CCCCAGAAGA CTATCCCTTC TTCAGAACTG TTAACATCGC TGACGGGAAG TGGCATCGGG 660
 TAGCAATCAG CGTGAGAGAG AAAACTGTGA CAATGATTGT TGATTGTAAG AAGAAAAACA 720
 CGAAACCACT TGATAGAAGT GAGAGAGCAA TTGTGTATAC CAATGGAATC ACGGTTTTTG 780
 GAACAAAGAT TTTGATGAA GAAGTTTTTG AGGGGGACAT TCAGCAGTTT TTGATCACAG 840
 GTGATCCCAA GGCAGCATAT GACTACTGTG AGCATTATAG TCCAGACTGT GACTCTTCAG 900
 CACCCAGGCG TGCTCAAGCT CAGGAACCTC AGATAGATGA GTATGCACCA GAGGATATAA 960
 TCGAATATGA CTATGAGTAT GGGGAAGCAG AGTATAAAGA GGCTGAAAGT GTAACAGAGG 1020
 GACCCACTGT AACTGAGGAG ACAATAGCAC AGACGGAGGC AAACATCGTT GATGATTTC 1080
 AAGAATACAA TTGTGATGAA ATGGAAAGTT ACCAGACAGA AGCTCCTAGG CATGTTTCTG 1140
 GGACAAATGA GCCAAATCCA GTTGAAGAAA TATTACTGA AGAATATCTA ACGGGAGAGG 1200
 ATTATGATTC CCAGAGGAAA AATTCTGAGG ATACACTATA TGAACAAACA GAAATAGAGC 1260
 GCAGGGATTG TGATCTCTG GTAGATGGAG AITTAGGGCA ATATGATTTT TATGAATATA 1320
 AAGAATATGA AGATAAACCA ACAAGCCCCC CTAAATGAAGA ATTTGGTCCA GGTGTACCA 1380
 CAGAAACGTA TATTACAGTA ACAAGCATAA ATGGCCATGG TGATATGGA GAGAAAGGAC 1440
 AGAAAGGAGA ACCAGCAGTG GTTGAGCCTG GTATGCTTGT CGAAGGAACA CCAGGACCA 1500
 CAGGACCTGC AGGTATTATG GGTCTCCAGG GTCTACAAGG CCCACTGGA CCCCCTGGTG 1560
 ACCCTGGCGA TAGGGGCCCC CCAGGACGTC CTGGCTTACC AGGGGCTGAT GGTCTACCTG 1620
 GTCTCTCTG TACTATGTTG ATGTTACCGT TCCGTTATGG TGGTGTATGT TCCAAAGGAC 1680
 CAACCATCTG TGCTCAGGAA GCTCAGGCTC AAGCTATTCT TCAGCAGGCT CGGATTGCTC 1740
 TGAGAGGCCC ACCTGGCCCA ATGGGTCTAA CTGGAAGACC AGTCTCTGTG GGGGGGCTG 1800
 GTTCATCTGG GGCCTAAGGT GAGAGTGGTG ATCCAGGTCC TCAGGGCCCT CGAGGGGTCC 1860
 AGGGTCCCCC TGGTCCAACT GGAAACCTG GAAAAGGGG TCGTCCAGT CAGATGGAG 1920
 GAAGAGGAAT GCGAGGAGAA CCTGGGGCAA AGGGAGATCG AGGGTTGAT GGACTTCCGG 1980
 GTCTGCTCC TGACAAAGGT CACAGGGGTG AACGAGGTCC TCAAGTCTCT CCAGGTCTCT 2040
 CTGGTGATGA TGGATATGAG GGAGAAGATG GAGAAATTGG ACCAAGAGGT CTTCCAGGTG 2100
 AAGCTGGCCC ACGAGTTTGT CTGGGTCCAA GGGGAACCTC AGGAGCTCCA GGGCAGCCTG 2160
 GTATGGCAGG GTATCTCTG CCCCCAGGAC CAAAAGGGAA CATGGTCCC CAAGGGGAGC 2220
 CTGGGCCCTC AGGTCAACAA GGGAAATCCA GACCTCAGGG TCTTCTCTGT CCACAGGTG 2280
 CAATTGTGTC TCTGTGTGAA AAAGGACCAC AAGGAAACC AGGACTTGCT GGACTTCTCT 2340
 GTGCTGATGG GCTCTCTGCT CATCTCTGGA AAGAAGGCCA GTCTGGAGAA AAGGGGGCTC 2400
 TGGGTCCCCC TGGTCCACAA GGTCTCTATT GATNNCCGGG CCCCCTGGGA GTAAAGGGAG 2460
 CAGATGTGTG CAGAGGTCTC AAGGGATCTA AAGGTGAAA GGGTGAAGAT GGTTTTCCAG 2520
 GATTCAAGAG TGACATGGGT CTAAAGAGTG ACAGAGGAGA AGTTGGTCAA ATTGGCCCAA 2580
 GAGGGAAGA TGGCCTTGAA GGACCCAAAG GTCGAGCAGG CCAACTGGA GACCCAGGTC 2640
 CTTAGGTCA AGCAGGAGAA AAGGGAAAAC TTGGAGTTCC AGGATTACCA GGATATCCAG 2700
 GAAGACAAGG TCCAAAGGCT TCCACTGGAT TCCTGGGTT TCCAGGTGCC AATGGAGAGA 2760
 AAGGTGCAAG GGGAGTAGCT GGCAAAACAG GCGCTCGGG TCAGGCTGCT CCAACGGGTC 2820
 CTCAGGTTTC AAGAGGTGCA AGAGGTCCCA CTGGGAAACC TGGGCCAAG GGCATCTCAG 2880
 GTGGCGATGG CCTCTCTGCT CCTCCAGGTG AAAGAGGTCC TCAAGACCT CAGGTCCTAG 2940
 TTGGATTCCC TGGACCAAAA GGCCCTCTCT GACCACAGG AAGGATGGG TGGCCAGGAC 3000
 ACCCTGGGCA ACGTGGGGAG ACTGGATTTC AAGGCAAGAC CGGCCCTCT GGGCCAGGG 3060
 GAGTGGTTGG ACCACAGGAG CCAACCGGTG AGACTGGTCC AATAGGGGAA CGTGGGTATC 3120
 CTGGTCTCTC TGGCCCTCTT GGTGAGCAAG GTCTCTCTGG TGCTGCAGGA AAAGAAGGTG 3180
 CAAAGGGTGA TCCAGGTCTT CAAGGTATCT CAGGGAAAAG TGGACCAACA GGATTACGTG 3240
 GTTTCCGAGG GGAAGAGAGT CTTCTCTGAG CTCAGGGTGC ACCTGGACTG AAAGGAGGGG 3300
 AAGGTCCCCA GGGGCCACCA GGTCCAGTTG GTCACACAG AGAACCTGGG TCAGCAGGTA 3360
 CAGCTGGCCC AATTGGTTTA CGAGGGGCCC CGGGACCTCA GGGTCTCTCT GGTCCAGCTG 3420
 GAGAGAAAGG TGCTCTCTGA GAAAAAGGTC CCCAAGGGCC TGCAGGGAGA GATGGAGTTC 3480
 AAGGTCTCTG TGGTCTCCCA GGGCCAGCTG GTCTCTCCGG CTCCTCTGGG GAAGACGGAG 3540
 ACAAGGGTGA AATTGGTGA CCGGGACAAA AAGGCAGCAA GGGTGGCAAG GGAGAAAATG 3600
 GCGCTCCCGG TCCCCAGGCT CTTCAAGGAC CAGTGGTGC CCTTGGAAIT GCTGGAGGTG 3660
 ATGGTGAACC AGGTCTTAGA GGACAGCAGG GGATGTTTGG GCAAAAAGGT GATGAGGGTG 3720
 CCAGAGGCTT CCTTGCACCT CTTGGTCCAA TAGGTCTTCA GGGTCTGCCA GGGCCACCTG 3780
 GTGAAAAGG TGAAAATGGG GATGTTGGTC CATGGGGGCC ACCTGTCTCT CCAGGCCCAA 3840

5
10
15
20
25
30
35
40

GAGGCCCTCA AGGTCCCAAT GGAGCTGATG GACCACAAGG ACCCCCAGGT TCTGTTGGTT 3900
 CAGTTGGTGG TGTGGGAGAA AAGGGTGAAC CTGGAGAAGC AGGAAACCCA GGGCCTCCTG 3960
 GGAAGACGAG TGTAGGCGGT CCCAAAGGAG AAGAGGAGGA GAAAGGGGAA GCTGGTCCAC 4020
 CTGGAGCTGC TGAACCTCCA GGTGCAAGG GGCCGCCAGG TGATGATGGC CCTAAGGGTA 4080
 ACCCGGGTCC TGTGTTGTTT CCTGGAGATC CTGGTCTCC TGGGGAACTT GGCCCTGCAG 4140
 GTCAAGATGG TGTGTTGGT GACAAGGGTG AAGATGGAGA TCTTGGTCAA CGGGTCTCTC 4200
 CTGGCCCATC TGGTAGAGCT GGCCCAACAG GTCTCTCTG AAAACGAGGT CCTCTGGAG 4260
 CTGCAAGTGC AGAGGGAGGA CAAGGTGAAA AAGGTGCTAA GGGGGAGCA GGTGCAGAA 4320
 GTCTCTCTGG AAAAACCCGC CCACTCGGTC CTGAGGAGC TGCAAGAAAG CCTGGTCCAG 4380
 AAGGTCTTCG GGGCATCCCT GGTCTGTGG GAGAACAAG TCTCCCTGGA GCTGCAGGCC 4440
 AAGATGGACC ACCTGTCTCT ATGGGACCTC CTGGCTTACC TGGTCTCAAA GGTGACCCTG 4500
 GCTCCAAGGG TGAAAGGGA CATCTGGTT TAATTGGCTT GATTGGTCT CCAGGAGAAC 4560
 AAGGGGAAA AGGTGACCGA GGGCTCCCTG GAATCAAGG ATCTCCAGGA GCAAAAGGGG 4620
 ATGGGGGAAT TCCTGGTCTT CTGGTCCCT TAGGTCCACC TGGTCTCCA GGTACCAG 4680
 GTCTCAAGG CCCAAGGGT AACAAAGGCT CTACTGGACC CGCTGGCCAG AAAGGTGACA 4740
 GTGGTCTTCC AGGGCTCTCT GGGCTCCAG GTCCACCTGG TGAAGTCATT CAGCTTTAC 4800
 CAATCTTGT CTCCAAAAA ACAGAGAAGC ATACTGAAG CATGCAAGCA GATGCAGATG 4860
 ATAATATTCT TGATTACTCG GATGGAATGG AAGAAATATT TGGTTCCTC AATTCCCTGA 4920
 AACCAAGCAT CGAGCATATG AAATTTCCAA TGGGTACTCA GACCAATCCA GCCCGAATT 4980
 GTAAAGACCT GCAACTCAGC CATCTGACT TCCAGATGG TGAATATTG ATTGATCTTA 5040
 ACCAAGGTTG CTGAGGAGAT TCCTTCAAAG TTTACTGTAA TTTACATCT GGTGGTGA 5100
 CTTGCATTTA TCCAGACAAA AAATCTGAGG GAGTAAGAAT TTCATCATG CCAAGGAGA 5160
 AACCAAGAA TGGTTTAGT GAATTTAAGA GGGGAAACCT GCTTTCATAC TTAGATGTTG 5220
 AAGGAAATTC CATCAATATG GTGCAATGA CATTCCTGAA ACTTCTGACT GCCTCTGCTC 5280
 GGCATAAATT CACCTACAC TGTCATCAGT CAGCAGCTG GTATGATGTG TCATCAGGAA 5340
 GTTATGACAA AGCACTTCGC TTCTCGGAT CAAATGATGA GGAGATGCTC TATGACAATA 5400
 ATCTTTTAT CAAACACTG TATGATGTT GTACGTCCAG AAAAGGCTAT GAAAAAAGT 5460
 TCATTGAAT CAATACACCA AAAATTGATC AAGTACCTAT TGTGATGTC ATGATCAGT 5520
 ACTTTGGTGA TCAGATCAG AAGTTCGGAT TTGAAGTTGG TCCTGTTTGT TTTCTGGCT 5580
 AAGATTAAAG CAAGAACAAT ATCAAAATCA CAGAAAAATG ACCTTGGTGC CACCAACCCA 5640
 TTTTGTGCCA CATGCAAGTT TTGAATAAGG ATGATGGA AACAACGCTG CATATACAGG 5700
 TACCATTAG GAATACCGA TGCCTTTGTG GGGGAGAAAT CACAGACAAA AGCTTTGAAA 5760
 ATCATAAAGA TATAAGTTGG TGTGGCTAAG ATGGAACAG GGCTGATTCT TGATTCCCAA 5820
 TTCTCACTC TCCTTTTCT ATTTGAATTT CTTTGGTGT GTAGAAAAA AAAAAAGAAA 5880
 AATATATATT CATAAAAAAT ATGGTGCTCA TTCTCATCCA TCCAGGATG ACTAAACAG 5940
 TGTGTTAAT AATTGTAAT TATTTGTGT ACAGTTCTAT ACTGTTATCT GTGTCCATT 6000
 CCAAAACTTG CAGSTGCCCT TGAATCCGC TGAATCTAAT TTATGAGGAT GCCGAACCT 6060
 GATGGCAATA ATATATGTAT TATGAAATG AAGTTATGAT TTCCGATGAC CCTAAGTCCC 6120
 TTTCTTGT TAATGATGAA ATCTCTTGT GTGTGTT

Seq ID NO: 361 Protein sequence
 Protein Accession #: NP_001845

45
50
55
60
65
70
75
80

1 11 21 31 41 51
 MEPSWRMT KRMLWDFVT TLALTFLFOA REVRGAAPVD VLKALDFHNS PEGISKITGF 60
 CTNRKNSKGS DTAYRVSQA QLSAPTQQLF PGGTFPEDFS ILFTVKPKKG IQSPLLSTYN 120
 EHGIIQIGVE VGRSPVFLFE DHTGKPAPEP YPLFRVTNIA DGKWRVAIS VEKKTVMIV 180
 DCKKKTTPKL DRSERAIVDT NGITVFGTRI LDEEVFEGDI QQFLITGDPK AAYDYCEHS 240
 PDCSSAPKA AQAGEPQIDE YAPEDIIEYD VEYGEAEYKE AESVTEGPTV TEETIAQTEA 300
 NIVDDPOEYN VGTMESYQTE APRHVSQNE PNPVERIFTE EYLTGEDYDS QRKNSIEDTLY 360
 ENKEIDGRDS DLLVDGDLGE YDFYKEYE DKPTSPPNEE FPGVPAETD ITETSINGHG 420
 AYGEKQKGE PAVVEPGMLV EGPPGAGPA GIMGPPGLQG FTGPPGDPGD RGPGRPGLP 480
 GADGLPGPPG TMLMLPRFYG GDSKSGPTIS AQEAQAQAIL QQARIALRGP PGPMGLTRP 540
 GPFVGGSSG AKGESGDPGP QGPRGVQGP GPTGKPKRG RFGADGGRGM PGEFGAKGDR 600
 GFDGLPLPG DKGRGERGP QGPPGPPDD GMRGEGEIG PRGLPGEAG RGLLGRGTP 660
 GAPQPGMAG VDGPPGPKGN MGPPGEGPP QQQGNPQPQ LPPGQGPIS PGEKGPQKP 720
 GLAGLPAGD PPGHPGKEGQ GSEKALGPP GPQGPFGXP PRGVKAGDV RGLKSGKEK 780
 GEDGPPGFG DMGLKDRGE VQIGPRGXG GPEGPKRAG PTGDPGSPQ AGEKGLGVP 840
 GLPGYGRG PKGSTGPPG PGANGKRGAR GVAGKPGPRG QRGPTGPRGS RGARGPTGK 900
 GPKGTSGDG PPGPPGERGP QGPPGPPGFP GPKGPPGPPG RMGCPGHPGQ RGETGPGKT 960
 GPPGPGVVG PQGPTGTGP IGERGYPGP GPPGQGLPG AAGKEGAKD PGPGISGKD 1020
 GPAGLPGFP ERGLPGAQGA PGLKGEGPQ GPPGVPSPG ERGSAGTAGP IGLRGRPGPQ 1080
 GPPGPAKEG APGEKGPQGP AGRDGVQGPV GLPGPAGPAG SPGEDGDKGE IGEPPGKGS 1140
 GKGEGPPG PPGGQGPVGA PGIAGDGEP GPRGQGMFG QKDEGARGF PGPPGPIGLQ 1200
 GLPGPPGEGK ENGDPVWPG PGPFGPRGP GPNAGDPQG PPGSVGVSG VGEKGEPEGE 1260
 GNPGPGEAG VGGPKGERGE KGEAGPPGA GPPGAKGPPG DDGPKNPGP VGPPGDPGP 1320
 GELGPAQDG VGGDKGEDG PQGPPGPPG GEAGPPGPPG KRGPAGAAGA EGRQGEKGA 1380
 GEAGGEPGP KTGVPVQGP AGKPGPEGLR GIPGVPGEQG LPGAAGDGP PGPMGPPGLP 1440
 GLKGDPSGK EKGHPGLIGL IGPPGEQGEK GDRGLPGTQG SPGAKGDDGI PGPAAGPLGP 1500
 GPPGLPGPQ PKNGKSTGP AQQKDSGLP GPPGPPGPPG EVIQLPLILS SKKTRRHTEG 1560
 MQADADDNIL DYSDMGEEIF GSLNSLKQDI EHKFPMTQ TNPARTCDL QLSHPDPPDG 1620
 EYWDPNQGC SGDSFKVYCN FTSGETCIY PDKSEGVRI SSWPKPKGS WFEFKRGKL 1680
 LSYLDVEGNS INMVQMTFLK LLTASARQNF TYHCHQSAW YDVSSGSYDK ALRFLGSNDE 1740
 EMSYDNNPFI KTLVDGCTSR KGYERTVIEI NTPKIDQVPI VDWMSDFGD QNQKFGFEVG 1800
 PFCPLG

Seq ID NO: 362 DNA sequence
 Nucleic Acid Accession #: NM_003107
 Coding sequence: 351-1775

85

1 11 21 31 41 51
 TTCCCCAGCA TTCGAGAAAC TCCTCTCTAC TTTAGCACGG TCTCCAGACT CAGCCGAGAG 60
 ACAGCAAACT GCAGCGCGGT GAGAGAGCGA GAGAGAGGGA GAGAGAGACT CTCCAGCCTG 120
 GGAATATAA CTCCTCTGCG AGAGGCGGAG AACTCTCTCC CCAATCTTT TGGGGACTTT 180

5	TCTCTCTTTA	CCCACCTCCG	CCCCTGCGAG	GAGTTGAGGG	GCCAGTTCGG	CCGCCGCGCG	240
	CGTCTTCCCG	TTCCGCGTGT	GCTTGGCCCG	GGGAACCGGG	AGGGCCCGGC	GATCGCGCGG	300
	CGGCCGCGCG	GAGGGTGTGA	GCGCGCGTGG	GCGCCGCGCG	AGCCGAGGCC	ATGGTGCAGC	360
	AAACCAACAA	TGCCGAGAAC	ACGGAAGGCG	TGCTGGCCGG	CGAGAGCTCG	GACTCGGGCG	420
	CGCGCTCGA	GCTGGGAATC	GCCTCCTCCC	CCACGCGCGG	CTCCACCGCC	TCCACGGGGG	480
	GCAAGGCGGA	CGACCCGAGC	TGGTGCAAGA	CCCCGAGTGG	GCACATCAAG	CGACCCATGA	540
	AGCCTTCAT	GGTGTGGTGG	CAGATCGAGC	GGCGCAAGAT	CATGGAGCAG	TGCGCCGAGA	600
	TGCAACAACG	CGAGATCTCC	AAGCGGCTGG	GCAAAACGCTG	GAAGCTGCTC	AAAGACAGCG	660
	ACAAGATCCC	TTTCTATTGA	GAGGCGGAGC	GGCTGCGCCT	CAAGCACATG	GCTGACTACC	720
10	CGACTACAA	GTACCGGCCC	AGGAAGAAGG	TGAAGTCCGG	CAACGCCAAC	TCCAGCTCCT	780
	CGGCCGCGCG	CTCCTCCCAAG	CCGGGGGAGA	AGGGAGACAA	GGTCGGTGGC	AGTGGCGGGG	840
	GGGCCGCGCG	GGCGCGCGCG	GGCGCGCGGA	GCAGCAACGC	GGGGGAGGGA	GGCGCGCGTG	900
	CGAGTGGCGG	CGGCGCCAAC	TCCAAACCGG	CGCAGAAAAA	GAGCTGCGGC	TCCAAAGTGG	960
	CGGGCGCGCG	GGCGGTGGG	GTTAGCAAAC	CGCACGCCAA	GCTCATCTGT	GCAGGCGGGG	1020
15	GGGGCGCGCG	GAAAGCAGCG	GCTGCCGCGG	CCGCTCCTT	CGCGCGGAA	CAGGCGGGGG	1080
	CGGCCGCGCT	GCTGCCCTTC	GGCGCGCGCG	CGACCCACCA	CTCGCTGTAC	AAGGCGCGGA	1140
	CTCCGAGCGC	CTCCGAGCTC	GCCTCCTCGG	CAGCCTCGGC	CTCGCGCGCC	CTCGCGCGCC	1200
	CGGGCAAGCA	CCTGGCGGAG	AAGAAGGTGA	AGCGCGTCTA	CCTGTTCCGC	GGCCTGGGCA	1260
	CGTCTGCTG	GCCCGTGGG	GGGTGGGCG	CGGAGCGGA	CCCCAGCGAC	CCCTGGGCG	1320
20	TGTACGAGGA	GGAGGGCGCG	GGCTGCTCGC	CCGACGCGCC	CAGCCTGAGC	GGCGCGCAGC	1380
	GGCGCGCGCT	GTCCCGCGCC	CGCGCGCGCT	CGCCCGCGGA	CCACGCGCGC	TACGCCAGCC	1440
	TGCGCGCGCG	CGCGCGCGCG	CGCTCCAGCG	CGCCCTCGCA	CGGCTCCTCC	TGCGCGCTCGT	1500
	CCCACTCCTC	CTCTTCTCTC	TOCTCGGCT	CCTCGTCTCT	CGACGACGAG	TTCCGAAGACG	1560
	ACCTGCTCGA	CCTGAACCCC	AGCTCAAACT	TTGAGAGCAT	GTCCCTGGGC	AGCTTCAGTT	1620
25	CGTCTGCGCG	GCTCGACCGG	GACCTGGATT	TTAACTTCGA	CGCCGGCTCC	GGCTCGCACT	1680
	TGAGATTCCC	GGACTACTGC	ACGCCGAGG	TGAGCGAGAT	GATCTCGGGA	GACTTGGCTCG	1740
	AGTCCAGCAT	CTCCAACCTG	GTTTTCACTT	ACTGAAGGGC	GCGCAGGCGA	GAGAGAAGGGC	1800
	CGGGGGGGGT	AGGAGAGGAG	AAAAAAAAG	TGAAAAAAG	AAACGAAAAG	GACAGACGAA	1860
	GAGTTTAAAG	AGAAAAGGGA	AAAAAGAAAG	AAAAAGTAAG	CAGGGCTCGT	TCCGCCCGCGT	1920
30	TCTGCTGCTG	GGATCAAGGA	GGCGCGCGCG	GTTTTGGACC	CGGCTCCCA	TCCCCACACT	1980
	TCCCGGGCGG	GGGAGCCACT	CTGCCACGCC	GGAGGGAGCG	GGAGGAGGAA	GAGGGTAGAC	2040
	AGGGGCGAG	TGTGATTGTT	GTTATTGATG	TTGTTGTTGA	TGGCAAAAAA	AAAAAGCGAC	2100
	TTGAGTTTGG	CTCCCTTTTG	CTTGAAGAGA	CCCCCTCCCC	CTTCCAACGA	GCTTCCGGAC	2160
	TTGCTCTGAC	CCCCAGCAG	AAGGCGAGTT	AGTTTTCTAG	AGACTTGAAG	GAGTCTCCCC	2220
35	CTTCTGCTG	CACCACTCTG	GTTTTGTTTT	ATTTTGCTTC	TTGGTCAAGA	AAGGAGGGGA	2280
	GAACCCAGCG	CACCCCTCCC	CCCCTTTTTT	TAAACGCGTG	ATGAAGACAG	AAGGCTCCCG	2340
	GGTGACGAAT	TTGGCGGATG	GCAGATGTTT	TGGGGGAACG	CGGGGACTGA	GAGACTCCAC	2400
	GCAGGCGAAT	TCCGTTTGG	GGCCTTTTTT	TCCTCCCTCT	TTTCCCTTGG	CCCCCTCTGC	2460
	AGCCGAGGGA	GGAGATGTTG	AGGGGAGGAG	GCCAGCCAGT	GTGACCGCGG	CTAGGAAATG	2520
40	ACCCGAGAAC	CCCGTTGGA	GCGCAGCAGC	GGGAGCTAGG	GCGCGGGGCG	GAGGAGGACA	2580
	CGAACTGGAA	GGGGGTTACG	GGTCAAACTG	AAATGGATTG	GCACGTTGGG	GAGCTGGCGG	2640
	CGGGCGCTGC	TGGGCTTCCG	CCTTCTTTTC	TACGTGAAT	CAGTGAGGTG	AGACTTCCCA	2700
	GACCCCGGAG	CGGTGGAGGA	GAGGAGACTG	TTTGATGTGG	TACAGGGGCA	GTCAGTGGAG	2760
45	GGCGAGTGGT	TTCCGAAAAA	AAAAAGAAAA	AAAAGGG			

Seq ID NO: 363 Protein sequence
Protein Accession #: NP_003098

50	1	11	21	31	41	51	
	MVQQTNAEN	TEALLAGESS	DSGAGLELGI	ASSFTPGSTA	STGGKADDPF	WCKTPSGHIK	60
	RPMNAFWWS	QIERKIMEQ	SPDMHNAEIS	KRLGKRKWL	KSDKIPFIR	EAERLRLKEM	120
	ADYPPDYKRP	RKKVKSNGAN	SSSSAAASSK	PGEKGDVKGG	SGGGHGGGG	GGSSNAGGG	180
55	GGGASGGGAG	SKPAGKSCG	SKVAGGAGGG	VSKPHAKLIL	AGGGGGKAA	AAAAASPAAE	240
	QAGAAALLPL	GAAADHESLY	KARTPSASAS	ASSAASASAA	LAAPGKELAE	KKVKRVYLP	300
	GLGTSSFPVG	GVGAGADPSD	PLGLYEEBGA	GCSPDAPSL	GRSSAASSFA	AGRSPADHRG	360
	YASLRASPA	PSSAPSHASS	SASSHSSSSS	SSGSSSDDE	FEDDLLDLNP	SSNFESMSLG	420
	SFSSSSALDR	DLDNFNPEPS	GSHFEPPDYC	TPEVSEMISG	DWLESSISNL	VFTY	

Seq ID NO: 364 DNA sequence
Nucleic Acid Accession #: U10860
Coding sequence: 123-2204

65	1	11	21	31	41	51	
	TGCCGCGTGC	TCCTCGACCA	GGCCTCCTTC	TCAACCTCAG	CCCGCGGCGC	CGACCCCTCC	60
	GGCACCCCTCC	CGCCCGCTCT	CGTACTGTGC	CCGTACCCGC	CGCGGCTCCG	GCCTTGGCCC	120
70	CGATGGCTCT	GTGCAACGGA	GACTCCAAGC	TGGAGAAATG	TGGAGGAGAC	CTTAAGGATG	180
	GCCACCAACCA	CTATGAAGGA	GCTGTGTGCA	TTCTGGATGC	TGGTGTCTCAG	TACGGGAAAG	240
	TCATAGACCG	AAGAGTGAGG	GAACTGTTCG	TGCAGTCTGA	AATTTTCCCC	TTGGAAACAC	300
	CAGCATTTGC	TATAAAGGAA	CAAGGATTCC	GTGCTATTAT	CATCTCTGGA	GSACCTAATT	360
	CTGTGTATGC	TGAAGATGCT	CCCTGGTTTG	ATCCAGCAAT	ATTCACTATT	GGCAAGCCTG	420
75	TTCTTGGAA	TTGCTATGGT	ATGCAGATGA	TGAATAAGGT	ATTGGAGGT	ACTGTGCACA	480
	AAAAAAGTGT	CAGAGAAGAT	GGAGTTTCA	ACATTAGTGT	GGATAATACA	TGTTTATTAT	540
	TCAGGGGCTC	TCGAAGGAA	GAAAGTTGTT	TGCTTACACA	TGGAGATAGT	GTAGACAAAG	600
	TAGCTGATGG	ATTCAAGGTT	GTGGCAGGTT	CTGGAACAT	AGTAGCAGGC	ATAGCAAAATG	660
	AATCTAAAAA	GTTATATGGA	GCACAGTTCC	ACCCTGAAGT	TGGCCTTACA	GAATAATGGA	720
80	AAGTAATACT	GAAGAATTTC	CTTTATGATA	TAGCTGGATG	CAGTGGAAAC	TTCAACCGTC	780
	AGAACAGAGA	ACTTGAGTGT	ATTGAGAGGA	TCAAAGAGAG	AGTAGGCACG	TCAAAAGTTT	840
	TGGTTTTACT	CAGTGGTGGG	GTAGACTCAA	CAGTTGTGAC	AGCTTTGCTA	AATGCTGCTT	900
	TGAACCAAGA	ACAAGTCATT	GCTGTGCACA	TTGATAATGG	CTTTATGAGA	AAACGAGAAA	960
	GGCAGTCTGT	TGAAGAGGCC	CTCAAAAAGC	TTGGAATTCA	GGTCAAAAGT	ATAAATGCTG	1020
85	CTCATTCTTT	CTACAATGGA	ACAACAACCC	TACCAATATC	AGATGAAGAT	AGAACCCAC	1080
	GGAAAAGAGT	TAGCAAAAAC	TTAAATATGA	CCACAAGTCC	TGAAGAGAAA	AGAAAATCAA	1140
	TTGGGGATAC	TTTTGTTAAG	ATTGCCAATG	AAGTAATTGG	AGAAATGAAC	TTGAAACCCAG	1200
	AGGAGGTTTT	CCTTGCCCAA	GGTACTTTAC	GGCCTGATCT	AATTGAAAGT	GCATCCCTTG	1260

5
 10
 15

```

TTGCAAGTGG CAAAGCTGAA CTCATCAAAA CCCATCAGAA TGACACAGAG CTCATCAGAA 1320
AGTTGAGAGA GGAGGGAAAA GTAATAGAAC CTCTGAAAAGA TTTTCATAAA GATGAAGTGA 1380
GAATTTTGGG CAGAGAACTT GGACTTCCAG AAGAGTITAGT TTCCAGGCAT CCATTTCCAG 1440
GTCCCTGGCT GGCATTCAGA GTAATATGTG CTGAAGAACC TTATATTGT AAGGACTTTC 1500
CTGAAACCAA CAATATTTTG AAAATAGTAG CTGATTTTTC TGCAAGTGTT AAAAAGCCAC 1560
ATACCCATT ACAGAGAGTC AAAGCCTGCA CAACAGAAGA GGATCAGGAG AAGCTGATGC 1620
AAATTACCA CTGCAATTCA CTGAATGCC TCTTGCTGCC AATTAAACT GTAGGTGTGC 1680
AGGGTACTAG TCGTTCCTAC AGTTACGTGT GTGGAATCTC CAGTAAAGAT GAACCTGACT 1740
GGGAATCACT TATTTTCTG GCTAGGCTTA TACCTCGCAT GTGTCAACAC GTTAACAGAG 1800
TTGTTTATAT ATTTGGCCCA CCAAGTAAAG AACCTCCTAC AGATGTTACT CCCACTTTCT 1860
TGACAACAGG GGTGCTCAGT ACTTTACGCC AAGCTGATTT TGAGGCCCAT AACATTCTCA 1920
GGGAGTCTGG GTATGCTGGG AAAATCAGCC AGATGCCGGT GATTTTGACA CCATTACATT 1980
TTGATCGGGA CCCACTTCAA AAGCAGCCTT CATGCCAGAG ATCTGTGGTT ATTCGAACCT 2040
TTATTACTAG TGACTTCATG ACTGGTATAC CTGCAACACC TGGCAATGAG ATCCCTGTAG 2100
AGGTGATATT AAAGATGGTC ACTGAGATTA AGAAGATTCC TGGTATTCTT CGAATTATGT 2160
ATGACTTAAC ATCAAAGCCC CCAGGAACCTA CTGAGTGGGA GTAATAAACT TC
  
```

Seq ID NO: 365 Protein sequence
 Protein Accession #: AAA60331

20
 25
 30
 35

```

1 11 21 31 41 51
MALCNGDSKL ENAGGDLKDG HHHYEGAVVI LDAGAQYGVK IDRRVRELTV QSEIFPLETP 60
APAIKEQGF AIIISGGPNS VYAEDAPWFD PAIFTIGKPV LGICYGMQMM NKVPGTGVHK 120
KSVREDDGVN ISVDNTECSLF RGLQKEEVLV LTHGSDVDKV ADGFKVVARV GNIVAGIANE 180
SKILYGAQFH PEVLGTENGK VILKNFLYDI AGCSGTFTVQ NRELECIREI KERVGTSKVL 240
VLLSGGVDSST VCTALNRLAL NQEQVIAVHI DNGFMKRKES QSVBEALKKL GIQVKVINA 300
HSPYNGTTLT PISDEBTRPR KRISKTLNMT TSPEEKRII GDTFVKIANE VIGEMNLKPE 360
EVFLAQGTLR PDLIESASLV ASGKAELIKT HNDTELIRK LBREKGVIEP LKDFHKDEVR 420
ILGRELGLPE ELVSRHPPFG PGLAIRVICA EEPYICKDFP ETNNILKIVA DFSASVKKPH 480
TLQVRVCACT TEEDQKLMQ ITSLSHSLNAP LLPKTVGVQ GDCRSYSYVC GISSKDEPDW 540
BSLIFLARI PRMCHNVNRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADFAHNILR 600
BSGYAGKISQ MPVILTPHMF DRDPLQKQPS QORSVVIRTP ITSDPMTGIP ATPGNEIPVE 660
VVLKMWTEIK KIPGISIRIMY DLTSKPPGTT EWE
  
```

Seq ID NO: 366 DNA sequence
 Nucleic Acid Accession #: NM_004219
 Coding sequence: 46-654

40
 45
 50
 55

```

1 11 21 31 41 51
GCGGCCTCAG ATGAATGCGG CTGTTAAGAC CTGCAATAAT CCAGAATGGC TACTCTGATC 60
TATGTTGATA AGGAAAATGG AGAACCAGGC ACCCGTGTGG TTGCTAAGGA TGGGCTGAAG 120
CTGGGCTCTG GACCTTCAAT CAAAGCCTTA GATGGGAGAT CTCAAGTTTC AACACCACT 180
TTTGGCAAAA CGTTCGATGC CCCACCAGCC TTACCTAAGG CTACTAGAAA GGCTTTGGGA 240
ACTGTCAACA GAGCTACAGA AAGTCTGTA AAGACCAAGG GACCCCTCAA ACAAACACAG 300
CCAAGCTTTT CTGCCAAAAA GATGACTGAG AAGACTGTTA AAGCAAAAAG CTCTGTTCCT 360
GCCTCAGATG ATGCTTATCC AGAAATAGAA AATTTCTTTC CCTTCAATCC TCTAGACTTT 420
GAGAGTTTGG ACCTGCTGTA AGAGCACCAG ATTGCGCACC TCCTCTGAG TGGAGTGCTC 480
CTCATGATCC TTGACGAGGA GAGAGAGCTT GAAAGCTGT TTGAGCTGGG CCCCCTTCA 540
CCTGTGAAGA TGCCCTCTCC ACCATGGGAA TCCAATCTGT TGCAGTCTCC TTCAAGCATT 600
CTGTGACACC TGGATGTGTA ATTGCCACCT GTTGTGCTGT ACATAGATAT TTAATTTTCT 660
TAGTGCTTCA GAGTTGTGT GTATTGTAT TAATAAGCA TTCTTCAACA GAAAAAATAA 720
AAAAAATA
  
```

Seq ID NO: 367 Protein sequence
 Protein Accession #: NP_004210

60
 65

```

1 11 21 31 41 51
MATLIYVDKE NGEPTGRVVA KDGLKLGSGP SIKALDGRSQ VSTPRFGKTF DAPPALPKAT 60
RKALGTVNRA TEKSVTKGPF LKQKQPSFSA KIMTEKTVKA KSSVPASDDA YPEIEKFPFP 120
NPLDFESFDL PEERQIAHLF LSGVPLMILD EERELEKLFP LGPPSPVKMP SPPWESNLQ 180
SPSSILSTLD VELPPVCCDI DI
  
```

Seq ID NO: 368 DNA sequence
 Nucleic Acid Accession #: NM_000597
 Coding sequence: 118-1104

70
 75
 80
 85

```

1 11 21 31 41 51
ATTGCGGGCG AGGGAGGAGG AAGAAGCGGA GGAGGCGGCT CCGGCTCGCA GGGCGGTGCA 60
CCTGCCCGCC CGCCCGCTCG CTCGCTCGCC CGCCCGCGCG CGCTGCCGAC CGCCAGCATG 120
CTGCCGAGAG TGGGCTGCCC CGCGCTCGCG CTGCCGCGCG CGCGCTGCTC CGCGCTGCTG 180
CCGCTGCTGC TGCTGCTACT GGGGCGGAGT GGGCGGGGCG CGCGGGCGCG CGCGGAGGTG 240
CTGTTCCGCT GCCCGCCCTG CACACCCGAG CGCTGCGCGC CCGCGCGGTT CCGCGCGGTT 300
CGCGCGCGCG CGCGGCTGCG CGCAGTGGCC GGAGGCGCCC GCATGCCATG CGCGGAGCTC 360
GTCCGCGGAG CGGGCTGCGG CTGCTGCTCG GTGTGCGCGC GGCTGGAGGG CGAGGCGTGC 420
GGCGTCTACA CCCCSCGCTG CGGCCAGGGG CTGCGCTGCT ATCCCCACCC GGGCTCCGAG 480
CTGCCCTGCG AGGCGCTGCT CATGGGCGAG GGCACTGTG AGAAGCGCGG GGACGCGGAG 540
TATGGGCGCA GCCCGGAGCA GGTTCGAGAG AATGGCGATG ACCACTCAGA AGGAGGCGCT 600
GTGGAGAACC ACCTGGACAG CACCATGAAC ATGTTGGGCG GGGGAGGCGG TGCTGGCCGG 660
AAGCCCTCTA AGTGGGTAT GAAGGAGCTG GCCGTGTTCC GGGAGAAGGT CACTGAGCAG 720
CACCGGCAGA TGGGCAAGGG TGGCAGCAT CACCTTGGCC TGGAGGAGCC CAAGAAGCTG 780
CGACCCACCC CTGCCAGGAC TCCCTGCCAA CAGGAACATG ACCAGGTCTT GGAGCGGATC 840
  
```

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TCCACCATGC GCCTTCCGGA TGAGCGGGGC CCTCTGGAGC ACCTCTACTC CCTGCACATC 900
 CCCAACTGTG ACNAGCATGG CCTGTACAAC CTCAAACAGT GCAAGATGTC TCTGAACGGG 960
 CAGCGTGGGG AGTGCTGGTG TGTGAACCCC AACACCGGGA AGCTGATCCA GGGAGCCCCC 1020
 ACCATCCGGG GGGACCCCGA GTGTCACTCT TTCTACAATG AGCAGCAGGA GGCTTGCGGG 1080
 GTGCACACCC AGCGGATGCA GTAGACCGCA GCCAGCCGGT GCCTGGGCGC CCTGCCCCCC 1140
 GCGCCTCTCC AAACACCGGC AGAAAACGGA GAGTGCTTGG GTGGTGGGTG CTGGAGGATT 1200
 TTCCAGTTCT GACACACGTA TTTATATTTG GAAAGAGACC AGCACCGAGC TCGGCACCTC 1260
 CCGGCGCTCT CTCTTCCAGC CTGCAGATGC CACACCTGCT CCTTCTTGCT TTCCCGGGGG 1320
 GAGGAAGGGG GTTGTGGTGG GGGAGCTGGG GTACAGGTTT GGGGAGGGGG AAGAGAAATT 1380
 TTTATTTTGG AACCCCTGTG TCCTTTTGC ATAAGATTAA AGGAAGGAAA AGT

Seq ID NO: 369 Protein sequence
 Protein Accession #: NP_000588

1 11 21 31 41 51
 MLPRVGCPL PLPPPPPLPL LPLLLLLLGA SGGGGGARAE VLPRCPPCTP ERLAACGPPP 60
 VAPPAVAAG AGGARMPCAE LVREPGCGCC SVCARLEGEA CGVYTPRCQG GLRCYPHPGS 120
 ELPLQALVMG EGTCEKRRDA EYGASFEQVA DNGDDHSEGG LVENHVDSTM NMLGGGGSAG 180
 RKPLKSGMKE LAVPREKYTE QHRQMGKGGK HHLGLEEPFK LRPPPARTPC QOELDQVLER 240
 ISTMRLPDER GPLEHLYSLR IPNCDKHGLY NLKQCKMSLN GQRGECWCVN PNTGKLIQGA 300
 PTIRGDPECH LFYNEQGEAC GVHTQRMQ

Seq ID NO: 370 DNA sequence
 Nucleic Acid Accession #: NM_004264
 Coding sequence: 6-440

1 11 21 31 41 51
 GGAACATGCG GGATCGGCTC ACGCAGCTTC AGGACGCTGT GAATTCGCTT GCAGATCAGT 60
 TTGTGAATGC CATTGGAGTA TTGCAGCAAT GTGGTCTCTC TGCCCTCTTC AATAATATTC 120
 AGACAGCAAT TAACAAGAGC CAGCCAGCTA ACCCTACAGA AGAGTATGCC CAGCTTTTTC 180
 CAGCACTGAT TGCAAGAACG GCAAAAGACA TTGATGTTTT GATAGATTCC TTACCCAGTG 240
 AAGAATCTAC AGCTCGCTTA CAGGCTGCTA GCTTGATATA GCTAGAAGAA GAAAACCATG 300
 AAGCTGCTAC ATGTGTGGAG GATGTTGTTT ATCGAGGAGA CATGCTCTCG GAGAAGATAC 360
 AAAGCGCACT TGCTGATATT GCACAGTCAC AGCTGAAGAC AAGAAGTGGT ACCCATAGCC 420
 AGTCTCTTCC AGACTCATAG CATCAGTGGG TACCATGTGG CTGAGAAAAG AACTGTTTGA 480
 GTGCCATTAA GAATCTGCA TCAGACTTAG ATACAAGCCT TACCAACAAAT TACAGAAACA 540
 TTAACACCTA TGACACATTA CCTTTTAGC TATTTTAAAT AGTCTTCTAT TTTCACCTCT 600
 GATAAGCTTA TAAATCATGA TTGAATCAGC TTTAAGCAT CATACCATCA TTTTTTAACT 660
 GAGTGAAATT ATTAAGGCAT GTAATACATT AATGAACATA ATATAAGGAA ACATATGTAA 720
 AATTCTGTGA TGACATAATT TATGCTCCA TTTTGTGTA TTGGCCAGTA CTTTACAAT 780
 C

Seq ID NO: 371 Protein sequence
 Protein Accession #: NP_004255

1 11 21 31 41 51
 MADRLTLQLD AVNSLADQFC NAIGVLQCG PPASFNNIQT AINKDQPANP TEEYAQLFAA 60
 LIARTAKDID VLIDSLPSEE STAAALQAASL YKLEENHEHA ATCVEDVVYR GDMLEKIQS 120
 ALADIAQSGL KTRSGTHSQS LPDS

Seq ID NO: 372 DNA sequence
 Nucleic Acid Accession #: AJ271091
 Coding sequence: 1-1113

1 11 21 31 41 51
 ATGGAGAATC AGGTGTTGAC GCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60
 CTGCGCGTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120
 CATTTCAGAG CTCAGGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180
 TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAACTGA CCCAGAGGCA GGTAAACATT 240
 ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAACCGACCA 300
 CTGTTTTTGG CTCTGACTT TGATCGTTGG CTGGATGAAT CTGATGCCGA AATGGAGCTC 360
 AGAGCTAAGG AAGAAGAGCG CCTAATATAA CTCCGACTGG AAAGCGAAGG CTCTCCTGAA 420
 ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGTCGA ATTCTTGGGA 480
 TTCTCCTGGA TCTTTGTCAA CCGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCCCTT 540
 TATGACACAT TCCATCTGT GGTGACATG ATGTATTCTT GCCAGATGCT GCCAGTTGTG 600
 GAAACTATCA ATGCAGCAAT TGGAGTCACT ACGTCACCGG TGCTGCCTTC TCTGATCCAG 660
 CTCTCTGAAA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGCAGAAC 720
 AAAGCTGTGG TTTTCTTTGT GTTTTATTG TGGAGTGCAA TTGAAATTTT CAGGTACTCT 780
 TTCTACATGC TGACGTGCAT TGACATGGAT TGGAGGTGCG TCACATGGCT TCGTTACACT 840
 CTGTGGATTG CCTTATATCC ACTGGGATGT TTGGCGGAAG CTGTCTCAGT GATTTCAGTC 900
 ATTCCAATAT TCAATGAGAC CGGACGATTG AGTTTCACAT TGCCATATCC AGTGAAAATC 960
 AAAGTTAGAT TTTCCTTTT TCTTCAGATT TATCTTATAA TGATATTTTT AGGTTTATAC 1020
 ATAAATTTTC GTCACTTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGGCT 1080
 CATGCCCTGT ATCCAGCGC TTTGGGAGGC TGA

Seq ID NO: 373 Protein sequence
 Protein Accession #: CAB69070

1 11 21 31 41 51
 MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60
 FLDLVKPEPV YKLTQRQVNI TVQKKVSQWW ERLTKQEKRP LPLAPDFDRW LDESDAEMEL 120
 RAKEERELNK LRLESEGSPE TLTNLRKGYL FMYNLVQFLG FSWIFVNLTV RFCILGKESF 180

YDFFHTVADM MYFCQMLAVV ETINAAIGVT TSPVLP SLIQ LLGRNFILFI IPGTMZEMQN 240
KAVVFPFVYL WSAIEIFRYS FYMLTCDMD WKVLTWLRYT LWIPLYPLGC LAEAVSVIQS 300
IPIFNETGRF SFTLPYPVKI KVRFSFFLQI YLIMIFLGLY INFRHLYKQR RLKMRAGAVA 360
HACDPSALGG

Seq ID NO: 374 DNA sequence
Nucleic Acid Accession #: NM_016395
Coding sequence: 1-1113

1 11 21 31 41 51
| | | | | |
ATGGAGAATC AGGTGTTGAC GCCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60
CTGCGCGTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120
CATTTCAAAG CTCAAGGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180
TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAAACTGA CCCAGAGGCA GGTAAACATT 240
ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300
CTGTTTTTGG CTCTGACTT TGATCGTTGG CTGGATGAAT CTGATGCGGA AATGGAGCTC 360
AGAGCTAAGG AAGAAGAGCG CCTAAATAAA CTCCGACTGG AAAGCGAAGS CTCTCTGAA 420
ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTGTGCA ATTCTTGGGA 480
TTCTCCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCCTTT 540
TATGACACAT TCCATACGTG GGCTGACATG ATGTATTCTT GCCAGATGCT GGCAGTTGTG 600
GAAACTATCA ATGCAGCAAT TGGAGTCACT ACCTCACCGG TGCTGCCTTC TCTGATCCAG 660
CTTCTTGAAA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGCAGAAC 720
AAAGCTGTGG TTTTCTTTGT GTTTTATTTG TGGAGTGCAA TTGAAATTTT CAGTACTCT 780
TTCTACATGC TGACGTGATC TGACATGGAT TGAAGGTGTC TCACATGGCT TCGTTACACT 840
CTGTGGGACT CCTTATATCC ACTGGGATGT TTGGCGGAAG CTGTCTCAGT GATTCACTCC 900
ATTCCAATAT TCAATGAGAC CGGACGATTC AGTTTCACAT TGCCATATCC AGTGAATATC 960
AAAGTTAGAT TTTCCTTTT TCTTCAGATT TATCTTATAA TGATATTTT AGGTTTATAC 1020
ATAAATTTTC GTCACTTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGGCT 1080
CATGCTGTG ATCCAGCGC TTTGGGAGGC TGA

Seq ID NO: 375 Protein sequence
Protein Accession #: NP_057479

1 11 21 31 41 51
| | | | | |
MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60
FLDLVKPEPV YKLTQRQVNI TVQRKVSQNM ERLTKQEKRP LFLAPDFDRW LDESDAEMEL 120
RAKEERLWLNK LRLESEGSPE TLINLRKGYL FMYNLVQFLG FSWIFVNLTV RFCILGKESF 180
YDFFHTVADM MYFCQMLAVV ETINAAIGVT TSPVLP SLIQ LLGRNFILFI IPGTMZEMQN 240
KAVVFPFVYL WSAIEIFRYS FYMLTCDMD WKVLTWLRYT LWIPLYPLGC LVEAVSVIQS 300
IPIFNETGRF SFTLPYPVKI KVRFSFFLQI YLIMIFLGLY INFRHLYKQR RRRYGRKRRR 360
STKKDLDDGF LPV

Seq ID NO: 376 DNA sequence
Nucleic Acid Accession #: NM_005987
Coding sequence: 1-270

1 11 21 31 41 51
| | | | | |
ATGAATTTCT AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA GCAGCAGCAG 60
GTGAAACACAC CTTGCCAGCC TCCACCCAGG GAACCATGCA TCCCAAAAC CAAGGAGCCC 120
TGCCAAACCA AGGTGCCCTGA GCCCTGCCAC CCCAAAGTGC CTGAGCCCTG CCAGCCCAAG 180
ATTCCAGAGC CTTGCCAGCC CAAGGTGCCCT GAGCCCTGCC CTTCAACGGT CACTCCAGCA 240
CCAGCCCAAG AGAAGACCAA GCAGAAGTAA

Seq ID NO: 377 Protein sequence
Protein Accession #: NP_005978

1 11 21 31 41 51
| | | | | |
MNSQQQKQPC TPPPQQQQQ VKQPCQPPQ EPCIPKTKEP CQPKVPEPCH PKVPEPQPK 60
IPEPCQPKVP EPCPSTVTPA PAQKTKQK

Seq ID NO: 378 DNA sequence
Nucleic Acid Accession #: NM_002105
Coding sequence: 74-505

1 11 21 31 41 51
| | | | | |
ACAGCAGTGA CACTGCGGCG GCGCTCTGTT CTAGTGTITG AGCGGTCTG CTTACCCGTT 60
CTACCTCGCT AGCATGTGCG GCGCGGCAAA GACTGGCGGC AAGGCCCGCG CCAAGGCCAA 120
GTCGCGCTCG TCGCGCGCGG GCCTCCAGTT CCCAGTGGGC CGTGACACC GGCTGCTGCG 180
GAAGGGCCAC TACGCGGAGC GCGTTGGCGC CGGCGCGCCA GTGTACCTGG CGGCACTGCT 240
GGAGTACCTC ACCGCTGAGA TCCTGGAGCT GCGCGGCAAT GCGGCCCGCG ACAACAAGAA 300
GACGCGAATC ATCCCCCGCC ACCTGCAGCT GGCCATCCGC AACGACGAGG AGCTCAACAA 360
GCTGCTGGGC GCGGTGACGA TCGCCCAAGG AGGCGTCTGT CCAACATCC AGGCGGTGCT 420
GCTGCCCAAG AAGACCAGCG CCACCGTGGG CGCGAAGCGC CCTCGGGCG GCAAGAAGGC 480
CACCCAGGCC TCCAGGAGT ACTAAGAGGG CCGCGCGCGC GCGCGGCCCG CCCAGCTCCC 540
CATGCCACCA CAAGGCCCT TTTAAGGGCC ACCACCGCCG TCATGGAAAG AGCTGAGCGG 600
CTTCAGACTG CGGCGCAAGC GGGCGCGCGC TCCCTTCCCC TCCCTCCCCC TCGCCCGCCT 660
TGGCGCGCGG GCTCGAGTC CCGCGCGCGC CCGCTCCCG TCCGACAGG CCGTCCCGCT 720
CGGCTTGGGC GCTCTTGGG GACCTCGGTG GCGCGGAAGA CCGGAGCCTG CCGGGGGGAG 840

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

```

GCCGCGCGCG CGCACCTGCG CGGCTCGGCG GTTCGTGACT CAGCGCGCCC ATCCCGAGTC 900
GCTAAGGGGCG TGCGGGGAGG CCGCAGCACC TTCTGSAAGA CTGGGCTTTC CGCTCTGACG 960
CAGGGCGGAG GTGGGCGAGTC CAGGCGGAGA GCCGCGGCGC CTGAAGGTGA GTGAGGCCCT 1020
CGGCGCTGCG AGCGGGGGTG TCTGGTACCC CCGCGGCGTG GTGCTTAGCC CAGGACTTTC 1080
AGACGCGCGC TGGCGGAGAG GCTTTGTGTG GAGAGACGCG ATGCGCGATT TCGGTCTGGC 1140
GCCCCCTCTG CGGCGGGGAC CCAGGCCCTT CACATCAGCT CTCCCTCCAT CTTTATTCTA 1200
AGGTCTGCGC TGCGGGCGGG ACGAAGCACT TGGTAACAGG CACATCTTCC TCCGAGTGA 1260
CTGCTCTCTA GGAGGACATT TAGGGGAGGG CAGAGGCGTG CAGTTTGGCT TCACGCGTGG 1320
CTATGTGGAC ACGAAGAGTC GTTTTGGGA ACGGACTGG CAGCCAGGCC TGTGGGGCCC 1380
COGAGCGCGC CCCATTTCCT TTCCAGCAA CTCACTCGG CAATCCAAGC ACCTAGATAC 1440
CAGCACAAGT CGGTTAATCC CTGTCTGGAC TGAGCCTCGG TTGGCTCTG AACTGGAATT 1500
CTGCAGCTAA CCCTTCCACG ACTAGAACCT TAGGCATTGG GGAGTTTATG ATGGACTAAT 1560
TTTATTAAGG GATTGTTTTT TTTT

```

Seq ID NO: 379 Protein sequence
Protein Accession #: NP_002096

1 11 21 31 41 51
1
20 MSGRGTGGK ARAKAKSRSS RAGLQFPVGR VHRLLRKHY AERVGAGAPV YLAAVLEYLT 60
AEILELAGNA ARDNKTRII PRHLQLAIRN DEELNKLGG VTIAQGGVLP NIQAVLLPKK 120
TSATVGPAP SGGKATQAS QEY

Seq ID NO: 380 DNA sequence
Nucleic Acid Accession #: AL136942
Coding sequence: 184-864

1 11 21 31 41 51
30
35
40
45
50
55
60
65
70
75
80
85

```

ACGCGTCCGG CAGAAGCTCG GAGCTCTCGG GGTATCGAGG AGGCAGGCC GCGGGGCGAC 60
GGCGGAGCGG GCGGGGAGCC GGAGCGGCGG AGGAGCGGCG AGCAGCGGCG CGGCGGGCTC 120
CAGGCGAGGC GGTGCAAGCT CCTGAAACT TGCGCGCGCG CTGCGGCCAC TGCGCGCGGA 180
GGATGAAGA TGGTGGCGCC CTGACGCGCG TTCTACTCCA ACAGCTGCTG CTTGTGCTGC 240
CATGTCCGCA CCGGCACCAT CCTGCTCGCG GTCTGTATC TGATCATCAA TGTGTGGTA 300
CTGTGATTT TATTGAGTGC CCTGGCTGAT CCGGATCAGT ATAACCTTTC AAGTCTCTGA 360
CTGGAGAGTG ACTTTGAGTT CATGGATGAT GCCAACATGT GCATTGCCAT TGGGATTCT 420
CTTCTCATGA TCCTGATATG TGCTATGGCT ACTTACGGAG CGTACAAGCA ACGCGCAGCC 480
TGGATCATCC CATTCCTCTG TTACAGATC TTGACTTTG CCTGGAACAT GTTGGTTGCA 540
ATCACTGTGC TTATTATATC AAACCTCCAT CAGGAATACA TACGGCAACT GCCTCCTAAT 600
TTTCCCTACA GAGATGATGT CATGTCACTG AATCCTACCT GTTTGGTCTT TATTATTCTT 660
CTGTATTATA GCAATTATCT GACTTTTAA GGTACTTGA TTAGCTGTGT TTGGAATGCG 720
TACCGATACA TCAATGGTAG GAACTCCTCT GATGTCTCTG TTTATGTTAC CAGCAATGAC 780
ACTAGCGTGC TGCTACCCCG GTATGATGAT GCCACTGTGA ATGGTGTCTG CAAGGAGCCA 840
CGCCACCATC ACCTGTCTCG CTAAGCCTTC AAGTGGGCGG AGCTGAGGCG AGCAGCTTGA 900
CTTTGAGAGC ATCTGAGCAA TAGTCTCTGT ATTTCACTTT TGCCATGAGC CTCCTCTGAGC 960
TGTGTTGTG CTGAATGCT ACTTTTAA ATTTAGATGT TAGATTGAAA ACTGTAGTTT 1020
TCAACATATG CTTTGCTAGA ACACGTGAT AGATTAACTG TAGAATTCTT CCTGTACGAT 1080
TGGGATATA ACGGGCTTCA CTAACCTTCC CTAGGCATTG AAACCTTCCC CAAATCTGAT 1140
GGACCTAGAA CTCTGCTTTT GTACCTGCTG GGCCCAAGAG TTGGGCATTT TTCTCTCTGT 1200
TCCTCTCTT TTGAAATGT AAAATAAAG CAAAATAGA CAACCTTTTC TTCAGCCTAT 1260
CCAGCATAGA GAACAAACCT TTATGGAAC AGGAATGTCA ATTTGTGAAT CATGTCTTA 1320
ATTAGGTAAA TAGAAGTCTT TATGTATGTG TTACAAGAT TTCCCCACA ACATCCTTTA 1380
TGACTGAAGT TCAATGACAG TTTGTGTTG GTGGTAAAGG ATTTTCTCCA TGGCCTGAAT 1440
TAAGACCAT AGAAGCACC AGGCGTGGG AGCAGTGACC ATCTACTGAC TGTCTCTGTG 1500
GATCTTGTGT CCAGGACAT GGGGTGACAT GCCTGTATG TGTAGAGGG TGGAAATGGAT 1560
GTGTTTGGCG CTGCATGGGA TCTGGTGGCC CTCTTCTCTT GGAATCACAT CCCACCCAG 1620
GGCCCGCTTT TACTAAGTGT TCTGCCCTAG ATTGGTTCAA GGAGGTCATC CAACTGACTT 1680
TATCAAGTGG AATTGGGATA TATTGATAT ACTTCTGCTT AACCAATGCT AAAAGGGTTT 1740
CTTTTCTCT GCAAGCTACA TCCTACTGCT TTGAACCTCC AAGTATGTCT AGTCACTCTT 1800
TAAATGTAA ACATTTTCAG AAAATGAGG ATTGCCTTCC TTGTATGCGC TTTTACCTT 1860
GACTACCTGA ATTGCAAGG ATTTTATAT ATTATATAT TACAAAGTCA GCAACTCTCC 1920
TGTTGGTTCA TTATGGAATG TGCTGTAAAT TAAGTCGTTT GCAATTAAAA CAAGGTTTGC 1980
CCACATCCAA AAAAAAAAAA AAAAA

```

Seq ID NO: 381 Protein sequence
Protein Accession #: CAB66876

1 11 21 31 41 51
70
75
80
85

```

MRMVPWTRF YSNSCLCCH VRTGILLGV WYLIINAVVL LILLSALADP DQYNFSSSEL 60
GGDFEFMDDA NMCIAIAISL LMILICAMAT YGAYKQRAAW IIPFCYQIF DFLNMLVAI 120
TVLIYPNSIQ EYIRQLPPNF PYRDDVMSVN FTCLVLIILL FISIILTFRG YLISCVWNCY 180
RYINGRNSSD VLVYVTSNDT TVLLPPYDDA TVNGAAKEPP PPVVA

```

Seq ID NO: 382 DNA sequence
Nucleic Acid Accession #: NM_002510
Coding sequence: 92-1774

1 11 21 31 41 51
80
85

```

CAGATGCCAG AAGAACACTG TTGCTCTTGG TGGACGGGCC CAGAGGAATT CAGAGTTAAA 60
CCTTGAGTGC CTGCTCCGCT GAGAATTCAG CATGGAATGT CTCTACTATT TCCTGGGATT 120
TCTGCTCTG GCTGCAAGAT TGCCACTTGA TGCCGCCAAA CGATTTCATG ATGTGCTGGG 180
CAATGAAGA CCTTCTGCTT ACATGAGGGA GCACAATCAA TTAAATGGCT GGTCTTCTGA 240
TGAAAATGAC TGGAAATGAA AACTCTACCC AGTGTGGAAG CGGGGAGACA TGAGGTGGAA 300
AAACTCTGG AAGGGAGGCC GTGTGCAGGC GGTCTGTACC AGTGAATCAC CAGCCCTCTG 360

```

	GGGCTCAAAT	ATAACATTTG	CGGTGAACCT	GATATTCCCT	AGATGCCAAA	AGGAAGATGC	420
	CAATGGCAAC	ATAGTCTATG	AGAAGAACTG	CAGAAATGAG	GCTGGTTTAT	CTGCTGATCC	480
	ATATGTTTAC	AACTGGACAG	CATGGTCAGA	GGACAGTGAC	GGGGAAAAAT	GCACCGGCCA	540
	AAGCCATCAT	AACGTCTTCC	CTGATGGGAA	ACCTTTTCTC	CACCACCCCG	GATGGAGAAG	600
5	ATGGAATTC	ATCTACGTCT	TCCACACACT	TGGTCAGTAT	TTCCAGAAAT	TGGGACGATG	660
	TTCACTGAGA	GTTCCTGTGA	ACACAGCCAA	TGTGACACTT	GGGCTTCAAC	TCATGGAAGT	720
	GACTGTCTAC	AGAAGACATG	GACGGGCATA	TGTTCCCATC	GCACAAGTGA	AAGATGTGTA	780
	CGTGGTAACA	GATCAGATTG	CTGTGTTTGT	GACTATGTTC	CAGAAGAACG	ATCGAAATTC	840
	ATCCGACGAA	ACCTTCCTCA	AAGATCTCCC	CATTATGTTT	GATGTCTCTGA	TTTATGATCC	900
10	TAGCCACTTC	CTCAATTATT	CTACCATTA	CTACAAGTGG	AGCTTCGGGG	ATAAATCTGG	960
	CCTGTGTTGT	TCCACCAATC	ATACTGTGAA	TCACAAGTAT	GTGCTCAATG	GAACCTTCAG	1020
	CCTTAACTTC	ACTGTGAAAG	CTGCAGCACC	AGGACCTTGT	CGCCACCCCG	CACCACCAAC	1080
	CAGACCTTCA	AAACCCACCC	CTTCTTTAGG	ACCTGTCTGT	GACAACCCCG	TGGAGCTGAG	1140
	TAGGATTCCT	GATGAAAACT	GCCAGATTAA	CAGATATGGC	CACCTTCAAG	CCACCATCAC	1200
15	AATTGTAGAG	GGAATCTTAG	AGGTTAATAT	CATCCAGATG	ACAGACGTCC	TGATGCGCGT	1260
	GCCATGGGCT	GAAAGCTCCC	TAATAGACTT	TGTCGTGACC	TGCCAAGGGA	GCATTCCCAC	1320
	GGAGGTCTGT	ACCATCATTT	CTGACCCACC	AGGACCTTGT	CGCCACCCCG	CACCACCAAC	1380
	CCCTGTGGAT	GTGGATGAGA	TGTGTCTGCT	GACTGTGAGA	CGAACCTTCA	ATGGGTCTGG	1440
	GACGTACTGT	GTGAACCTCA	CCCTGGGGGA	TGACACAAGC	CTGGCTCTCA	CGAGCACCGT	1500
20	GATTTCTGTT	CCTGACAGAG	ACCCAGCCTC	GCCTTTAAGG	ATGGCAAAAC	GTGCCCTGAT	1560
	CTCCGTTGGC	TGCTTGGCCA	TATTTGTCA	TGTGATCTCC	CTCTTGGTGT	ACAAAAAACA	1620
	CAAGGAATAC	AAACCAATAG	AAATAGTCTC	TGGGAATGTG	GTCAAGAGCA	AAGGCTTGAG	1680
	TGTCTTCTTC	AACCGTGCAA	AAGCCGTGTT	CTTCCCGGGA	AACCAAGAAA	AGGATCCGCT	1740
25	ACTCAAAAAC	CAAGAATTTA	AAGGAGTTTC	TTAAATTTTG	ACCTTGTGTT	TGAAGCTCAC	1800
	TTTTCACTGC	CATTGATGTG	AGATGTGCTG	GAGTGGCTAT	TAACCTTTTT	TTCTTAAAGA	1860
	TTATTGTAA	ATAGATATTG	TGGTTTGGGG	AAGTTGAATT	TTTTATAGGT	TAAATGTCTAT	1920
	TTTAGAGATG	GGGAGAGGGA	TTATACTGCA	GGCAGCTTCA	GCCATGTTGT	GAAGCTGATA	1980
	AAAGCAACTT	AGCAAGGCTT	CTTTTCATTA	TTTTTTATGT	TTCACTTATA	AAGTCTTAGG	2040
30	TAACATAGAG	GATAGAAACA	CTGTGTCCCG	AGAGTAAGGA	GAGAAGCTAC	TATTGATTAG	2100
	AGCCTAAACC	AGCTTAACTG	CAAGAAGAGG	CGGATACTTT	TCAGCTTTCC	ATGTAACCTG	2160
	ATGCATAAAG	CCAATGTAGT	CCAGTTTCTA	AGATCATGTT	CCAAGCTAAC	TGAATCCAC	2220
	TTCAATACAC	ACTCATGAAC	TCCTGATGGA	ACAAATACAG	GCCCAAGCCT	GTGGTATGAT	2280
	GTGCACACTT	GCTAGACTCA	GAAGAAATAC	TACTCTCATA	AATGGGTGGG	AGTATTTTGG	2340
	TGACAACCTA	CTTTGCTTGG	CTGAGTGAAG	GAATGATATT	CATATATTCA	TTTATTCAT	2400
35	GGACATTTAG	TTAGTGCCTT	TTATATACCA	GGCATGATGC	TGAGTGACAC	TCTTGTGTAT	2460
	ATTTCCAAAT	TTTTGTATAG	TCGCTGCACA	TATTTGAAAT	CATATATTAA	GACTTTCCAA	2520
	AGATGAGGTC	CTGTGTTTTT	CATGGCAACT	TGATCAGTAA	GGATTTTACC	TCTGTTTGT	2580
	ACTAAACCA	TCTACTATAT	GTTAGACATG	ACATTCTTTT	TCTCTCTTTC	CTGAAAAATA	2640
	AAGTGTGGGA	AGAGACAAAA	AAAAAATA				

Seq ID NO: 383 Protein sequence
Protein Accession #: NP_002501

	1	11	21	31	41	51	
45	MECLYYFLGP	LLLAARLPD	AAKRFHDVLG	NERPSAYMRE	HNQLNGWSSD	ENDWNEKLYP	60
	VWRKGRDMRWK	NSKKGGRRVQA	VLTSDSPALV	GSNITFAVNL	IFPRQCKEDA	NGNIVYEKNC	120
	RNEAGLSADP	YVYNWTAWSE	DSGGENGTGQ	SHNVFPDQK	PPPHHPGWR	WNFIYVPHL	180
	GQYFQKLGRS	SVRVSVNTAN	VTLPQLMVE	TVYRRHGRAY	VPIAQVKDQV	VVTDQIPVFN	240
50	TMFQKNDNRS	SDETFLKDLF	IMFDVLHDP	SHPLNYSTIN	YKWSFGDNTG	LFVSTNHTVN	300
	HTVVLNGTFS	LNLTVKAAAP	GPCPPPPPPP	RPSKPTPSLG	PAGDNPLELS	RIPDENQIN	360
	RYGHFQATIT	IVEGILEVNI	IQMTDVLMPV	PWPESSLIDF	VVTQGSIPIT	EVCTIISDPT	420
	CEITQNTVCS	PVDVDEMCLL	TVRRTFNGSG	TYCVNLTLGD	DTSLALTSTL	ISVPRDRPAS	480
	PLRMANSLI	SVGLAIPVT	VISLLVYKKE	KEYNPIENSP	GNVVRSKGLS	VFLNRAKAVP	540
55	FPNQEKDPL	LKNQEFKGV					

Seq ID NO: 384 DNA sequence
Nucleic Acid Accession #: NM_001134
Coding sequence: 48-1877

	1	11	21	31	41	51	
60	TCCATATTGT	GCITCCACCA	CTGCCAATAA	CAAAATAACT	AGCAACCATG	AAGTGGGTGG	60
	AATCAATTTT	TTTAATTTTC	CTACTAAAT	TTACTGAATC	CAGAACACTG	CATAGAAATG	120
65	AATATGGAAT	AGCTTCCATA	TTGGATTCTT	ACCAATGTAC	TGCAGAGATA	AGTTTACCTG	180
	ACCTGGCTAC	CATATTTTTT	GCCAGTTTGT	TTCAAGAAAG	CACCTACAAG	GAAGTAAGCA	240
	AAATGGTGAA	AGATGCATTG	ACTGCAATTG	AGAAACCCAC	TGGAGATGAA	CAGTCTTCAG	300
	GGTGTTTAGA	AAACCAGCTA	CCTGCCTTTC	TGGAAGAACT	TTGCCATGAG	AAAGAAATTT	360
70	TGGAGAAGTA	CGGACATTCA	GACTGCTGCA	GCCAAAGTGA	AGAGGGAAGA	CATAACTGTT	420
	TTCTTGCACA	CAAAAGCCCC	ACTCCAGCAT	CGATCCCACT	TTTCCAAGTT	CCAGAACCTG	480
	TCACAAGCTG	TGAAGCATAT	GAAGAAGACA	GGGAGACATT	CATGAACAAA	TTTCTTTATG	540
	AGATAGCAAG	AAGGCATCCC	TTCTGTATG	CACCTACAAT	TCTTCTTTGG	GCTGCTCGCT	600
	ATGACAAAAT	AATTTCCATC	TGCTGCAAG	CTGAAATATG	AGTTGAATGC	TTCCAAACAA	660
	AGGCAGCAAC	AGTTACAAAA	GAATTAAGAG	AAAGCAGCTT	GTTAAATCAA	CATGCATGTG	720
75	CAGTAATGAA	AAATTTTGGG	ACCCGAACCT	TCCAAGCCAT	AACGTGTACT	AAACTGAGTC	780
	AGAAGTTTAC	CAAGTTAAT	TTTACTGAAA	TCCAGAAACT	AGTCCTGGAT	GTGGCCCATG	840
	TACATGAGCA	CTGTTGCAGA	GGAGATGTGC	TGGATTGTCT	GCAGGATGGG	GAAGAAATCA	900
	TGTCCTACAT	ATGTTCTCAA	CARGACACTC	TGTCAACAAA	AATAACAGAA	TGCTGCAAAC	960
	TGACCAAGCT	GGAGCGTGGT	CAATGTATTA	TTTCAATGAA	AAATGATGAA	AAACCTGAAG	1020
80	GTCTATCTCC	AAATCTAAAC	AGGTTTITAG	GAGATAGAGA	TTTAAACCAA	TTTCTTCTAG	1080
	GGGAAAAAAA	TATCTTCTTG	GCAAGTTTGT	TTTCAATGAA	TTCAAGAAAG	CATCTCTCAG	1140
	TTGCTGTCTC	AGTAATTTCT	AGAGTTGTCT	AAGGATACCA	GGAGTTATTT	GAGAAGTGTG	1200
	TCAGAGCTGA	AAACCTCTCT	GAATGCCAAG	ATAAAGGAGA	AGAGAAATTA	CAGAAATACA	1260
	TCCAGGAGAG	CCAGCATTG	GCAAGCGGAA	GCTGCGGCTC	CTTCCAGAAA	CTAGAGAGAT	1320
85	ATTACTTACA	AAATGCGTIT	CTGTTGCTT	ACACAAAGAA	AGCCCCCCAG	CTGACCTCGT	1380
	CGGAGCTGAT	GGCCATCACC	AGAAAAATGG	CAGCCACAGC	AGCCACTTGT	TGCCACTTCA	1440
	GTGAGGACAA	ACTATTGGCC	TGTGGCGAGG	GAGCGGCTGA	CATTATTATC	GGACACTTAT	1500

GTATCAGACA TGAATGACT CCAGTAAACC CTGSGTGTGG CCAGTGCTGC ACTTCTTCAT 1560
 ATGCCAACAG GAGGCCATGC TTCAGCAGCT TGGTGGTGGG TGAAACATAT GTCCCTCCTG 1620
 CATTCTCTGA TGACAAGTTC ATTTTCCATA AGGATCTGTG CCAAGCTCAG GGTGTAGCGC 1680
 TGCAAAACGAT GAAGCAAGAG TTTCTCATTG ACCTTGTGAA GCAAAAGCCA CAAATAACAG 1740
 AGGAACAAC TGAAGCTGTC ATTGCAGATT TCTCAGGCTT GTTGGAGAAA TGCTGCCAAG 1800
 GCCAGGAACA GGAAGTCTGC TTTGCTGAAG AGGGACAAAA ACTGATTTC AAAAACTCGT 1860
 CTGCTTTGGG AGTTTAAATT ACTTCAGGGG AAGAGAAGAC AAAACGAGTC TTTTATTCTG 1920
 TGTGAACCTT TCTCTTAAAT TTTAACTGAT TTAACACTTT TTGTGAATTA ATGAAATGAT 1980
 AAAGACTTTT ATGTGAGATT TCCTTATCAC AGAAATAAAA TATCTCCAAA TG

Seq ID NO: 385 Protein sequence

Protein Accession #: NP_001125

1 11 21 31 41 51
 MKWVESIFLI FLNFTESRT LHRNEYGIAS ILDSYQCTAE ISLADLATIF FAQFVQEATY 60
 KEVSKMVIDA LTAIEKPTGD EQSSGCLENQ LPAFLLEELCH EKEILEKYGH SDCCSQSEEG 120
 RENCFLAHKK PTPASIPLFQ VPBPVTSCEA YEEDRETFMN KFIYEIARRH PFLYAPTILL 180
 WAARYDKIIP SOCKAENAVE CPQTKAATVT KELRESSLLN QHACAVMKNF GTRTFQAITV 240
 TKLSQKPTKV NPTBIQKLVL DVAHVHEHCC RGDVLDCLQD GEKIMSYICS QODTSLNKIT 300
 ECKKLTTLER GQCIHAEND EKPEGLSPNL NRFLGDRDFN QFSSGKKNIF LASFVHEYSR 360
 RHPQLAVSGV LRVAQGYQEL LEKCFQTEHP LECQDKGEEE LQKYIQESQA LAKRSGGLFQ 420
 KLGEYLLQNA FLVAYTKKAP QLTSSSELMAI TRKMAATAAT CCQLSEDKLL ACGEGAADI 480
 IGHLCIRHEM TPVNPVGVC CTSSYANRRP CFSLSVVDET YVPPAFSDDK FIFHKDLQQA 540
 QGVVALQTMKQ EFLINLVKQK PQITEQLEA VIADFSGLLE KCCQGGQEV CFAEEGQKLI 600
 SKTRALGV

Seq ID NO: 386 DNA sequence

Nucleic Acid Accession #: NM_002205.1

Coding sequence: 1..3149

1 11 21 31 41 51
 ATGGGGAGCC GGACGCCAGA GTCCCCCTCTC CACGCGGTGC AGCTGCGCTG GGGCCCCCGG 60
 CGCCGACCCC CGCTSSSTGCC GCTGCTGTGT CTGCTSSSTGC CGCGGCCACC CAGGGTCGGG 120
 GGCTTCAACT TAGACGCGGA GGGCCGAGCA GTACTCTCGG GGGCCCCGGG CTCTCTCTTC 180
 GGATTCTCAG TGGAGTTTGA CCGCCGCGGA ACAGACGGGG TCAGTGTGCT GGTGGGAGCA 240
 CCCAAGGCTA ATACGAGCCA GCCAGGAGTG CTGCAGGGTG GTGCTGTCTA CCTCTGTCTT 300
 TGGGGTGCCA GCGCCACACA GTGCACCCCTC ATTGAATTTG ACAGCAAAAG CTCTCGGCTC 360
 CTGGAGTCCCT CACTGTCCAG CTCAGAGGGA GAGGAGCCTG TGGAGTACAA GTCCTTGCGA 420
 TGGTTCGGGG CAACAGTTCG AGCCCATGGC TCCTCCATCT TGGCATGCGC TCCACTGTAC 480
 AGCTGGCGCA CAGAGAAGGA GCCACTGAGC GACCCCGTGG GCACCTGCTA CCTCTCCACA 540
 GATAACTTCA CCCGAATTCT GAGATATGCA CCCTGCCGCT CAGATTTCAG CTGGGCAGCA 600
 GGACAGGGTT ACTGCCAAGG AGGCTTCAGT GCCGAGTTCA CCAAGACTGG CCGTGTGGTT 660
 TTAGGTGGAC CAGGAAGCTA TTTCTGGCAA GGCCAGATCC TGTCTGCCAC TCAGGAGCAG 720
 ATTGCAAGAT CTATTATCCC CGAGTACCTG ATCAACCTGG TTCAGGGGCA GCTGCAGACT 780
 CGCCAGGCCA GTTCCATCTA TGATGACAGC TACCTAGGAT ACTCTGTGGC TGTGGTGA 840
 TTCAGTGGTG ATGACACAGA AGACTTTGTT GCTGGTGTGC CCAAGGGGAA CCTCACTTAC 900
 GGCTATGTCA CCATCTCTAA TGCTCAGAC ATTGCATCCC TCTACAATT CTCAGGGGAA 960
 CAGATGGGCT CCTACTTTGG CTATGCAAGT GCGGCCACAG ACGTCAATGG GGACGGGCTG 1020
 GATGACTTGC TGGTGGGGGC ACCCTGCTC ATGGATCGGA CCCCTGACGG GCGGCTCTAG 1080
 GAGGTGGGCA GGGTCTACGT CTACCTGCAG CACCCAGCGG GCATAGAGCC CACGCCAC 1140
 CTTACCCCTA CTGGCCATGA TGAGTTTGGC GATTTTGGCA GCTCCTTGAC CCCCCTGGGG 1200
 GACCTGGACC AGGATGGCTA CAATGATGTG GCCATCGGGG CTCCTTTTGG TGGGAGAGCC 1260
 CAGCAGGGCT TAGTGTTTGT ATTTCTGGGG GGGCCAGGAG GGGTGGGCTC TAAGCCTTCC 1320
 CAGGTCTCTG AGCCCTGTGG GGCAGCCAGC CACACCCAG ACTTCTTTGG CTCTGCCCTT 1380
 CGAGGAGGCC GAGACCTGGA TGGCAATGGA TATCCTGATC TGAATTTGGG GTCCTTTGGT 1440
 GTGGACAAGG CTGTGGTATA CAGGGGCGCG CCCATCGTGT CCGCTAGTGC CTCCTTCA 1500
 ATCTTCCCGG CCATGTTCAA CCCAGAGGAG CGGAGCTGCA GCTTAGAGGG GAACCTCTGT 1560
 GCGCTGCATCA ACCCTAGCTT CTGCTCAAT GCTTCTGGAA AACACGTTGC TGACTCCATT 1620
 GGTTTCACAG TGAACCTTCA GCTGGACTGG CAGAAGCAGA AGGGAGGGGT ACGGCGGGCA 1680
 CTGTCTCTGG CCTCCAGGCA GGCACCCCTG ACCCAGACCC TGCTCATCCA GAATGGGGCT 1740
 CGAGAGGATT GCAGAGAGAT GAAGATCTAC CTCAGGAAGC AGTCAGAATT TCGAGACAAA 1800
 CTCTCGCCGA TTCACATCGC TCTCAACTTC TCCTTGGACC CCCAAGCCCC AGTGGACAGC 1860
 CACGGCTCTA GCGCCAGCCC ACATTATCAG AGCAAGAGCC GGATAGAGGA CAAGGCTCAG 1920
 ATCTTGCTGG ACTGTGGAGA AGACAACATC TGTGTGCTGT ACCTGCAGCT GGAAGTGT 1980
 GGGGAGCAGA ACCATGTGTA CCTGGGTGAC AAGAATGCCG TGAACCTCAC TTTCCATGCC 2040
 CAGAATGTGG GTGAGGGTGG CGCTATGAG GCTGAGCTTC GGGTCACCGC CCTCCAGAG 2100
 GCTGAGTACT CAGGACTCGT CAGACACCCA GGGAACTTCT CCAGCCTGAG CTGTGACTAC 2160
 TTTGCGGTGA ACCAGAGCCG CTTGCTGGTG TGTGACCTGG GCAACCCCAT GAAGGCAGGA 2220
 GCCAGTCTGT GGGGTGGCCT TCGGTTTACA GTCCCTCATC TCGGGACAC TAAGAAAACC 2280
 ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTCGCAAG CGAGTGGTT 2340
 TCCTTTCCGG TCTCCGTGGA GGCTCAGGCC CAGGTACCCC TGAACGTTGT CTCCAAGCCT 2400
 GAGGCAGTGC TATTCCAGT AAGCGACTGG CATCCCCGAG ACCAGCCTCA GAAGGAGGAG 2460
 GACCTGGGAC CTGCTGTCCA CCATGTCTAT GAGCTCATCA ACCAAGGCCC CAGCTCCATT 2520
 AGCCAGGGTG TGTGGAACCT CAGCTGTCCC CAGGCTCTGG AAGGTGAGCA GCTCCTATAT 2580
 GTGACCAGAG TTAAGGGACT CAACTGCACC ACCAATCACC CCATTAAACC AAAGGGCCTG 2640
 GAGTGGATC CCGAGGGTTC CCGCACCCAC CAGCAAAAAC GGGGAAGCTCC AAGCCGCGAG 2700
 TCTGCTTCTT CGGAGCCTCA GATCCTGAAA TGCCCGGAGG CTGAGTGT 2760
 TGTGAGCTG GGCCTCTGCA CCAACAGAG AGCCAAAGTC TGCAGTTGCA TTTCCAGTCA 2820
 TGGGCCAAGA CTTTCTTGCA GCGGGAGCAC CAGCCATTTA GCCTGCAAGT TGAGGCTGTG 2880
 TACAAGCCCC TGAAGATGCC CTACCGAATC CTGCTCTGGC AGCTGCCCCA AAAAGAGCGT 2940
 CAGGTGGGCA CAGCTGTGCA ATGGACCAAG GCAGAAGGCA GCTATGGCGT CCCACTGTGG 3000
 ATCATCATCC TAGCCATCCT GTTTGGCCTC CTGCTCTAG GTCTACTCAT CTACATCCTC 3060
 TACAAGCTTG GATTCTTCAA ACGCTCCCTC CCATATGGCA CCGCATGGA AAAAGCTCAG 3120
 CTCAGCCCTC CAGCCACCTC TGATGCTGTA

Seq ID NO: 387 Protein sequence
Protein Accession #: NP_002196.1

```

5      1      11      21      31      41      51
|      |      |      |      |      |
MGRSTPESPL HAVQLRWGPR RRPPLLPLLL LLLPPPPRVG GFNLDAEAPA VLSGPPGSFF 60
GPSVEFYRPG TDGVSVLVGA PKANTSQPGV LQGGAVYLCF WGASPTQCTP IEFDSKGSRL 120
LESSLSSEG EEPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPVGTCTYLT 180
10 DNFTRILEYA PCRSDFSWAA GQGYCQGGPS AEFTKTGRVV LGGPGSYFWQ GQILSATQEQ 240
IAESYYPEYL INLVQQQLQT RQASSIYDDG YLGYSVAVGE FSGDDTEDEV AGVPKGNLTY 300
GVVTILNGSD IRSLYNFSGE QMASYPGYAV AATDVNGDGL DDLVVGAPLL MDRTPDGRPQ 360
EVRGVVYVLO HPAGIEPTPT LTLTGHDFFG RFGSSLTPLG DLDQDGYNDV AIGAPPGGET 420
QQGVVVFVPG GPGGLGSKPS QVLQPLWAAS HTDFFPGSAL RGRDLGNG YPDLIVGSFG 480
15 VDKAVVYRGR PIVSASASLT IFPAMFNPEE RSCSLGNFV ACINLSFCLN ASGKHVADSI 540
GPTVLEQLDW QKQKGVVRRR LPLASRQATL TQTLIIQNGA REDCREMKIY LRNESEFRDK 600
LSPIHIALNF SLDPQAPVDS HGLRPLAHYQ SKSRIEDKAQ ILLDCGEDNI CVPDLQLEVF 660
GBQNVHVLGD KNALNLTFAH QNVGEGGAYE AELRVTAPE AEYSGLVRRH GNFSSLSCDY 720
FAVNSRLRLV CDLGNFMKAG ASLWGLRFT VPHLRDTKCT IQDFQLLSK NLNNSQSDVV 780
20 SFRLSVEAQA QVTINGVSKP EAVLFPVSDW HPRDQPKKEE DLGPAVHVY ELINQGPSSI 840
SQGVLELSCP QALEGQQLLY VTRVTGLNCT TNHFINKGL ELDFEGSLHH QOKREAPSR 900
SASSGFQILK CPBAECFRLR CELGPLHQOE SQSLQLHFRV WAKTFQREH QPFSLQCEAV 960
YKALKMPYRI LPRQLPKER QVATAVQWTK AEGSYGVPLW IILAILFGL LLLGLLIYIL 1020
YKLGPFKRSL PYGTAMEKAQ LKPPATSDA

```

Seq ID NO: 388 DNA sequence
Nucleic Acid Accession #: NM_002425
Coding sequence: 26..1453

```

30      1      11      21      31      41      51
|      |      |      |      |      |
AAAGAAGGTA AGGGCAGTGA GAATGATGCA TCTTGCAATC CTTGTGCTGT TGTGTCTGCC 60
AGTCTGCTCT GCCTATCCTC TGAGTGGGCG AGCAAAAGAG GAGGACTCCA ACAAGGATCT 120
TGCCAGCAAA TACCTAGAAA AGTACTACAA CCTCGAAAAG GATGTGAAAC AGTTTAGAAG 180
35 AAAGACAGT AATCTCATTT TTAATAAAT CCAAGGAATG CAGAAGTTCC TTGGGTTGGA 240
GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGGAGT 300
TCTGACGTTT GGTCACTTCA GCTCCTTTCC TGGCATGCGG AAGTGGAGGA AAACCCACCT 360
TACATACAGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTG ATTCTGCCAT 420
TGAGAAAGCT CTGAAAGTCT GGAAGAGGCT GACTCCACTC ACATTCTCCA GGCTGTATGA 480
AGGAGAGGCT GATATAATGA TCTCTTTCG AGTTAAAGAA CATGGAGACT TTTACTCTTT 540
TGATGGCCCA GGACACAGTT TGGCTCATGC CTACCCACCT GGACCTGGGC TTTATGGAGA 600
TATTCACTTT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATTCCT 660
CGTTGCTGCT CATGAACCTG GCCACTCCCT GGGGCTCTTT CACTCAGCCA ACACCTGAAGC 720
TTTGATGTAC CCACTCTACA ACTCATTAC AGAGCTCGCC CAGTTCGGCC TTTGCAAGA 780
45 TGATGTGAAT AGGATTCAGT CTCTACCGG ACCTCCCGCT GCCTCTACTG AGGAACCCCT 840
GGTGCCCAAA AATCTGTGTC CTTCGGGATC TGAGATGCCA GCCAAGTGTG ATCTGCTTT 900
GTCTCTCGAT GGCATCAGCA CTCTGAGGGG AGAATATCTG TTCTTTAAG ACAGATATT 960
TTGGGCAAGA TCCACTGGA ACCCTGAACC TGAATTTCTT TTGATTTCTG CATTTTGCC 1020
CTCTCTTCCA TCATATTGGG ATGCTGCATA TGAAGTTAAC AGCAGGACA CCGTTTTTAT 1080
50 TTTTAAAGGA AATGAGTTCT GGGCCATCAG AGGAAATGAG GTACAAGCAG GTATCCAAG 1140
AGGCATCCAT ACCTCGGTTT TTCTCCCAAC CATAAGGAAA ATTGATGCAG CTGTTCTGA 1200
CAAGGAAAAG AAGAAAACAT ACTTCTTTC AGCGGACAAA TACTGGAGAT TTGATGAAA 1260
TAGCCAGTCC ATGGAGCAAG GCTTCCCTAG ACTAATAGCT GATGACTTTC CAGGAGTTGA 1320
GCCTAAGGTT GATGCTGTAT TACAGGCATT TGGATTTTTC TACTTCTTCA GTGGATCATC 1380
55 ACAGTTTGA TTTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAGA GTAAACAGTG 1440
GTTACATTGC TAGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTAAAT AAATCTAATA 1500
ATTATTCTAT TAATGTATTA TGAGCCAAAA TGGTTAATTT TTCTGTCATG TCTGTGACT 1560
GAAGAAGATG AGCCTTGCAG ATATCTGCAT GTGTCATGAA GAATGTTTCT GGAATCTTTC 1620
ACTTGTCTTT GAATTCGACT GAACAGAAAT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680
60 ATGTATTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGGCCTGTTC 1740
CTT

```

Seq ID NO: 389 Protein sequence
Protein Accession #: NP_002416

```

65      1      11      21      31      41      51
|      |      |      |      |      |
MHLAFVLVLC LPVCSAYPLS GAAKEEDSNK DLAQQYLEKY YNLEKDVKQF RRKDSNLIVK 60
KIQGMQKFLG LEVTGKLDTD TLEVMRKPRC GVPDVGHFSS PFGMPKWRKT HLTIRIVNYT 120
70 PDLPRDAVDS AIEKALKVNE EVTPLTF SRL YEGEADIMIS FAVKEHGDY SFDGPGHSLA 180
HAYPPGGLGY GDHIFDDDEK WTEDASGTNL FLVAAHELGH SLGLFHSANT EALMYPLVNS 240
FTELAQFRLS QDDVNGIQSL YGPPFPASTE PLVPTKSVPS GSEMPAKCDP ALSFDAISTL 300
RGEYLFKKDR YFWRRSHWNP EPEFHLSAF WPSLPSYLD AYEVSNSRDTV FIFRGNFPA 360
75 IRGNEVQAGY PRGIHTLGFP PTIRKIDAAV SDKEKKTYF PAADKYWRPD ENSQSMEQGF 420
PRLIADDFPG VEPKVDVAVLQ AFGFFYFSSG SSQFEFDPNA RMVTHILKSN SWLHC

```

Seq ID NO: 390 DNA sequence
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

```

80      1      11      21      31      41      51
|      |      |      |      |      |
ATGCACAGCT TTCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTG ACACAGCTTC 60
CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120
85 TACTACAACC TGAAGAAATG TGGGAGGCAA GTTGAAAGC GGAGAAATAG TGGCCAGTG 180
GTTGAAAAAT TGAAGCAAA GCAGGAATTC TTTGGGCTGA AAGTACTGG GAAACAGAT 240
GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCGTGATG GCCTCAGTTT 300

```

5 GTCCTCACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAAT 360
 TACAGGCCAG ATTGGCCAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCTCACTC 420
 TGGAGTAATG TCACACCTCT GACATTCAAC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480
 ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540
 CTTGCTCATG CTTTTCACCC AGGCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600
 GAAAGGTGGA CCAACAATTT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGAACCTC 660
 GGCCATTCTC TTGGATTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCTAGCTAC 720
 ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATG ATGGCATCCA AGCCATATAT 780
 10 GGAGCTTCCC AAAATCCTGT CCAGCCCATC GGCCCAAAA CCCCAAAAGC ATGTGACAGT 840
 AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900
 TTCTACATGC GCACAAATCC CTCTACCCG GAAGTTGAGC TCAATTTTCA TTCTGTTTTT 960
 TGGCCACAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020
 CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
 15 CCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140
 CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT 1200
 GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTTCT 1260
 GGAATTGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTCTA TTTCTTTCAT 1320
 GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
 AATAGCTGGT TCAACTGCAG GAAAAATTAG

Seq ID NO: 391 Protein sequence
Protein Accession #: NP_002412.1

25 1 11 21 31 41 51
 MHSFPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
 VEKLKQMQEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTGEGNPRWE QTHLTYRIEN 120
 YTPDLPRADV DHAIEKAFQL WSNVTPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180
 LAHAFQPGPG IGGDAHFDED ERWTNNFREY NLHRVAAHAL GHSLGLSHST DIGALMYPST 240
 30 TFSGDVQLAQ DDIDGIAIY GRSQNPVQPI GPQTPKACDS KLTFDAITTI RGEVMPFKDR 300
 FYMRINFPYP EVELNFIQVF WPQLPNGLEA AYEAFADRDEV RPFKGNKYWA VQQQNVLHGY 360
 PKDIYSSPGF PRTVKHIDAA LSEENTGKTY FFVANKYWRV DEYKRSMDPG YPKMIAHDFP 420
 GIGHKVDVAF MKDGFYFFPH GTRQYKFDPK TKRILTLQKA NSWFNCRKN

Seq ID NO: 392 DNA sequence
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

40 1 11 21 31 41 51
 ATGCACAGCT TTCCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTG ACACAGCTTC 60
 CCAGCGACTC TAGAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120
 TACTACAACC TGAAGAAATG TGGGAGGCAA GTTGAAGAAC GGAGAAATAG TGGCCCATG 180
 45 GTTGAATAAT TGAAGCAATG GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCATG 240
 GCTGAACACC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCCTGATG GGCTCAGTTT 300
 GTCCCTCACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAAT 360
 TACAGGCCAG ATTGGCCAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCTCACTC 420
 TGGAGTAATG TCACACCTCT GACATTCAAC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480
 ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540
 50 CTTGCTCATG CTTTTCACCC AGGCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600
 GAAAGGTGGA CCAACAATTT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGCCCTC 660
 GGCCATTCTC TTGGATTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCTAGCTAC 720
 ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATG ATGGCATCCA AGCCATATAT 780
 GGAGCTTCCC AAAATCCTGT CCAGCCCATC GGCCCAAAA CCCCAAAAGC ATGTGACAGT 840
 55 AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900
 TTCTACATGC GCACAAATCC CTCTACCCG GAAGTTGAGC TCAATTTTCA TTCTGTTTTT 960
 TGGCCACAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020
 CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
 60 CCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140
 CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT 1200
 GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTTCT 1260
 GGAATTGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTCTA TTTCTTTCAT 1320
 GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
 AATAGCTGGT TCAACTGCAG GAAAAATTAG

Seq ID NO: 393 Protein sequence
Protein Accession #: NP_002412.1

70 1 11 21 31 41 51
 MHSFPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
 VEKLKQMQEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTGEGNPRWE QTHLTYRIEN 120
 YTPDLPRADV DHAIEKAFQL WSNVTPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180
 75 LAHAFQPGPG IGGDAHFDED ERWTNNFREY NLHRVAAHAL GHSLGLSHST DIGALMYPST 240
 TFSGDVQLAQ DDIDGIAIY GRSQNPVQPI GPQTPKACDS KLTFDAITTI RGEVMPFKDR 300
 FYMRINFPYP EVELNFIQVF WPQLPNGLEA AYEAFADRDEV RPFKGNKYWA VQQQNVLHGY 360
 PKDIYSSPGF PRTVKHIDAA LSEENTGKTY FFVANKYWRV DEYKRSMDPG YPKMIAHDFP 420
 GIGHKVDVAF MKDGFYFFPH GTRQYKFDPK TKRILTLQKA NSWFNCRKN

Seq ID NO: 394 DNA sequence
Nucleic Acid Accession #: NM_014331.2
Coding sequence: 1..1506

1 11 21 31 41 51

	ATGGTCAGAA	AGCCTGTTGT	GTCCACCATC	TCCAAAGGAG	GTTACCTGCA	GGGAAATGTT	60
	AACGGGAGGC	TGCCCTCCCT	GGGCAACAAG	GAGCCACCTG	GGCAGGAGAA	AGTGCAGCTG	120
	AAGAGGAAAG	TCACTTTACT	GAGGGGAGTC	TCCATTATCA	TTGGCACCAT	CATTGGAGCA	180
5	GGAACTCTTA	TCTCTCTTAA	GGGCGTGCTC	CAGAACAACG	GCAGCGTGGG	CATGTCTCTG	240
	ACCATCTGGA	CGGTGTGTGG	GGTCTGTGCA	CTATTTGGAG	CTTTGTCTTA	TGCTGAATTG	300
	GGAAACAATA	TAAAGAAATC	TGGAGGTCAT	TACACATATA	TTTTGGAAAT	CTTTGGTCCA	360
	TTACAGAGCT	TTGTACGAGT	CTGGGTGGAA	CTCCTCATAA	TACGCCCTGC	AGCTACTGCT	420
	GTGATATCCC	TGGCATTGGG	ACGCTACATT	CTGGAAACAT	TTTTTATTCA	ATGTGAAATC	480
10	CCTGAACCTG	CGATCAAGCT	CATTACAGCT	GTGGGCATAA	CTGTAGTGAT	GGTCTCTAAAT	540
	AGCATGAGTG	TCAGCTGGAG	CGCCCGGATC	CAGATTTTCT	TAACCTTTTG	CAAGCTCACA	600
	GCAATTTCTA	TAATTATAGT	CCCTGGAGTT	ATGCAGCTAA	TAAAGGTCA	AACGCAGAAC	660
	TTTAAAGAGC	CGTTTTCAGG	AAGAGATTCA	AGTATTACGC	GGTTGCCACT	GGCTTTTAT	720
	TATGGAATGT	ATGCATATGC	TGGCTGGTTT	TACCTCAACT	TTGTTACTGA	AGAAGTAGAA	780
15	AACCCGAA	AAACCATTC	CCTTGCAATA	TGTATATCCA	TGGCCATTGT	CACCATTTGGC	840
	TATGTGCTGA	CAATGTGGC	CTACTTTACG	ACCATTAAATG	CTGAGGAGCT	GCTGCTTTCA	900
	AATGCAGTGG	CAGTGACCTT	TTCTGAGCGG	CTACTGGGAA	ATTTCTCATT	AGCAGTTCCG	960
	ATCTTTGTTG	CCCTCTCCTG	CTTTGGCTCC	ATGAACGGTG	GTGTGTTTGC	TGTCTCCAGG	1020
	TTATTCTATG	TTGGCTCTCG	AGAGGGTCAC	CTTCCAGAAA	TCCTCTCCAT	GATTCTATGC	1080
20	CGCAAGCACA	CTCCTCTACC	AGCTGTTATT	GTTTTGCACC	CTTTGACAA	GATAATGCTC	1140
	TTCTCTGGAG	ACCTCGACAG	TCCTTTGAAT	TTCTCTCAGT	TTGCCAGGTG	GCTTTTATT	1200
	GGGCTGGCAG	TTGCTGGGCT	GATTTATCTT	CGATACAAAT	GCCAGATAT	GCATCGTCTC	1260
	TTCAAGGTGC	CAGTTTCTCA	CCGAGCTTTG	TTTTCTCTCA	CATGCTCTCT	CATGGTTGCC	1320
	CTTTCCCTCT	ATTCCGAGCC	ATTAGTACA	GGGATTGGCT	TGTCATCAC	TCTGACTGGA	1380
25	GTCCCTGGGT	ATTATCTCTT	TATTATATGG	GACAAGAAAC	CCAGGTGGTT	TAGAATAATG	1440
	TCAGAGAAAA	TAACAGAAAC	ATTACAAATA	ATACTGGAAG	TGTACCAGA	AGAAGATAAG	1500
	TTATGAACCTA	ATGGACTTGA	GATCTTGGCA	ATCTGCCCAA	GGGAGACAC	AAAAAGGGA	1560
	TTTTTACTTC	ATTTTCTGAA	AGTCTAGAGA	ATTACAACTT	TGGTGATAAA	CAAAAGGAGT	1620
	CAGTTATTTT	TATTCATATA	TTTTAGCATA	TTGGAACATA	TTTCTAAGAA	ATTTAGTAT	1680
30	AACTCTATGT	AGTTATAGAA	AGTGAATATG	CAGTTATTCT	ATGAGTCCCA	CAATTCTTGA	1740
	GTCTCTGATA	CCTACTTATT	GGGTTTAGGA	GAAAAGACTA	GACAAATTACT	ATGTGGTCAT	1800
	TCCTTACAA	ATATGTTAGC	ACGGCAAGAA	ACCTTCAAT	TGAAGACTGA	GATTTTCTG	1860
	TATATATGGG	TTTTGTAAAG	ATGGTTTAC	ACACTACAGA	TGCTTACT	GTGAAAAGTG	1920
	TTTTCAATTC	TGAAAAAAG	CATACATCAT	GATTATGGCA	AAGAGGAGAG	AAAGAAATTT	1980
35	ATTTTACATT	GACATTGCAT	TGCTTCCCCT	TAGATACCAA	TTTAGATAAC	AAACACTCAT	2040
	GCTTTAATGG	ATTATACCCA	GAGCACTTTG	AACAAAGGTC	AGTGGGGATT	GTGGAATACA	2100
	TTAAAGAAGA	GTTTCTAGGG	GCTACTGTTT	ATGAGACACA	TCCAGGAGTT	ATGTTTAAAT	2160
	AAAAATCCTT	GAGAATTTAT	TATGTACAGT	GTTTTTCAT	TCATTATCAG	GAAAGTTTAG	2220
	TTATCTGTCA	TTTTTTTTTT	TCACATCAGT	TTGATCAGGA	AAGTGTATAA	CACATCTTAG	2280
40	AGCAAGAGTT	AGTTTGGTAT	TAAATCCTCA	TTAGAACAAC	CACCTGTTTC	ACTAATAACT	2340
	TACCCCTGAT	GAGTCTATCT	AAACATATGC	ATTTTAAAGC	TTCAAATTAC	ATTATCAACA	2400
	TGAGAGAAAT	AACCAACAAA	GAAGATGTTT	AAAAATATAG	TCCCATATCT	GTAATCATAT	2460
	CTACATGCCA	TGTTAGTAA	TCGAAGTTT	TTTAAATTTA	TGGCTATTTT	TACACGATGA	2520
	TGAATTTTGA	CAGTTTGTGC	ATTTTCTTTA	TACATTTTAT	ATTCCTCTGT	TAAAATATCT	2580
45	CTTCAGATGA	AACCTGTCAG	ATTAATTAGG	AAAAGGCATA	TATTAACATA	AAAAATGCCA	2640
	AAGAAATGTC	GCTGTAAATA	AGATTTACAA	CTGATGTTTC	TAGAAAATTT	CCACTTCTAT	2700
	ATCTAGGCTT	TGTCTAGTAA	TTCCACACCT	TAATTATCAT	TCAACTTGCA	AAAGAGACAA	2760
	CTGATAAGAA	GAAAATTGAA	ATGAGATCT	GTGGATAAGT	GTTTGTGTTT	AGAAGATGTT	2820
	GTTTGTCCAG	TATTAGAAAA	TACTGTGAGC	CGGGCATGGT	GGCTTACATC	TGTAATCCCA	2880
50	GCATTTGGG	AGGCTGAGGG	GGTGGATCAC	CTGAGGTGGG	GAGTCTAGTA	CCAGCCTGAC	2940
	CAACATGGAG	AAACCCCATC	TCTACTAAAA	ATACAAAATT	AGCTGGGCAT	GGTGGCCAT	3000
	GCTGTGTAAT	TCAGCTATGT	AGGAGGCTGA	GGCAGGAGAA	TTGCTTGAAC	CCGGAGGCGG	3060
	GAGGTTGCAG	TGAGCCAAGA	TTGCACCCT	GTACTCCAGC	CTGGGTGACA	AAGTCAGACT	3120
	CCATCTCCAA	AAAAAAAAAA	AAAA				

Seq ID NO: 395 Protein sequence
Protein Accession #: NP_055146.1

	1	11	21	31	41	51	
60	MVRKPVVSTI	SRGGYLQGNV	NGRLPSLGNK	EPFGQEKVQL	KRKVTLRGRV	SIIIGTIIGA	60
	GIPIGPKGVL	QNTGVSVMSL	TIWTVCGVLS	LFGALSYAEL	GTTIKKSGGH	YTYILEVFGP	120
	LPAPVRVWVE	LLIIRPAATA	VISLAFGRYI	LEPFFIQCEI	PELAIKLITA	VGITVVMVLN	180
	SMSVWSARI	QIPLTFCKLT	AIIIIIVPGV	MQLIKGQTQN	PKDAPSGRDS	SITRLPLAFY	240
65	YGMAYAGWP	YLNFTVEEVE	NPEKTIPLAI	CISMAITIGV	YVLTNVAYFT	TINAEELLLS	300
	NAVAVTFSER	LLGNFSLAVP	IFVALSCFGS	MNGGVFAVSR	LPYVASREGH	LPEILSMIHW	360
	RKHTPLPAVI	VLEPLTMIML	PSGLDLSLLN	FLSFARWLPF	GLAVAGLIYL	RYKCPDMHRP	420
	SEKIRTRLQI	ILEVVFEEBK	L				480

Seq ID NO: 396 DNA sequence
Nucleic Acid Accession #: NM_006528
Coding sequence: 57..764

	1	11	21	31	41	51	
75	GCCGCCAGCG	GCTTCTCGG	ACGCCCTGCC	CAGCGGGCCG	CCCGACCCCC	TGCACCATGG	60
	ACCCGCTCG	CCCCCTGGG	CTGTGATTC	TGCTGCTTTT	CCTGACGGAG	GCTGCACCTG	120
	GGCATGTGTC	TCAGAGGCCA	ACAGGAAATA	ACCGGAGAT	CTGTCTCCTG	CCCCTAGACT	180
80	ACGACCCCTG	CCGGGCCCTA	CTTCTCCGTT	ACTACTACGA	CAGGTACACG	CAGAGCTGCC	240
	GCCAGTTCTT	GTACGGGGGC	TGCGAGGGCA	ACGCCAACAA	TTTCTACACC	TGGGAGGCTT	300
	GCGACGATGC	TTGCTGGAGG	ATAGAAAAAG	TTCCCAAGT	TTGCCGGCTG	CAAGTGAGTG	360
	TGGACGAGCA	GTGTGAGGGG	TCCACAGAAA	AGTATTTCTT	TAATCTAAGT	TCCATGACAT	420
	GTGAAAAATT	CTTTTCCGGT	GGGTGTCAAC	GGAAACCGAT	TGAGAACAGG	TTTCCAGATG	480
85	AAGCTACTTG	TATGGGCTTC	TGCGCACCAA	AGAAAAATCC	ATCATTTTGC	TACAGTCCAA	540
	AAGATGAGGG	ACTGTGCTCT	GCCAAATGTA	CTCGCTATTA	TTTAAATCCA	AGATACAGAA	600
	CCTGTGATGC	TTTACCTAT	ACTGGCTGTG	GAGGGAATGA	CAATAACTTT	GTTAGCAGGG	660

AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGA AAAAGAA AAGAAGATG CCAAAGCTTC 720
GCTTTGCCAG TAGAATCCGG AAAATTCGGA AGAAGCAATT TTAACATTC TTAATATGTC 780
ATCTTGTTTG TCITTTATGCG TTAITTCGCT TTATGGTGT ATCTGAAGAA TAATATGACA 840
GCAATGAGAA ACAATCATT GGTGATTAT TCACCAGTTT TTATTAATAC AAGTCACTTT 900
TTCAAAATTT TGGATTTTT TATATATAAC TAGCTGCTAT TCAATGTGA GTCTACCAT 960
TTTAATTAT GGTCAACTG TTTGTGAGAC GAATTCCTGC AATGCATAAG ATATAAAAGC 1020
AAATATGACT CACTCATTTT TTGGGTCGT ATTCTGTATT TCAGAAGAGG ATCATAACTG 1080
AAACAACATA AGACAATATA ATCATGTGCT TTTAATCAT TTAGAATAA AAAGGACTAG 1140
CC

Seq ID NO: 397 Protein sequence
Protein Accession #: NP_006519

1 11 21 31 41 51
MDPARPLGLS ILLFLFLTEAA LGDAAQEPTG NNAEICLLPL DYGPCRALLL RYYYDRYTQS 60
CRQFLYGGCE GNANNFYTWB ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFNLSSM 120
TCEKFPSSGCG HRNRIENRFP DEATCMGFCA PKKIPSFCSY PKDEGLCSAN VTRYFENPRY 180
RTCDAPFTYTG CGGNDNNFVS REDCIRACAK ALKKKKKMPK LRFASRIIRKI RKKQF

Seq ID NO: 398 DNA sequence
Nucleic Acid Accession #: NM_001508.1
Coding sequence: 1..1361

1 11 21 31 41 51
ATGGCTTCAC CCAGCCTCCC GGGCAGTGAC TGCTCCCAAA TCATTGATCA CAGTCATGTC 60
CCCGAGTTTG AGGTGGCCAC CTGGATCAAA ATCACCCTTA TTCTGGTGA CTGATCATC 120
TTCTGTATGG GCCTTCCTGG GAACAGCGTC ACCATTGGG TCACCAGGT GCTGCAGAAG 180
AAAGGATACT TGCAGAAGGA GGTGACAGAC CACATGGTGA GTTTGGCTTG CTGGACATC 240
TTGGTGTTC TCATCGGCAT GCCATGGAG TTCTACAGCA TCATCTGGAA TCCCTGACC 300
ACGTCCAGCT CACCCCTGTC CTGCAAGCTG CACACTTTCC TCTTCGAGGC CTGCAGTAC 360
GCTACGCTGC TGCACGTGCT GACGCTCAGC TTTGAGOGCT ACATCGCCAT CTGTCAACCC 420
TTCAAGTACA AGCTGTGTGC GGGACCTTGC CAGGTGAAGC TGCTGATTGG CTTCGTCTGC 480
GTCACCTCCG CCCTGGTGGC ACTGCCCTTG CTGTTTGCCA TGGGTACTGA GTACCCCTCTG 540
GTGAACGTGC CCAGCCACCG GGGTCTCACT TGCAACCGCT CCAGCACCCG CCACCACGAG 600
CAGCCCGAGA CTCCAATAT GTCCATCTGT ACCAACCTCT CCAGCCGCTG GACCGTGTTC 660
CAGTCCAGCA TCTTCGGCGC CTTCGTGGTC TACCTCGTGG TCCTGCTCTC CTGAGCCTTC 720
ATGTGCTTGA ACATGATGCA GGTGCTCATG AAAAGCCAGA AGGGCTCGCT GGCCGGGGGC 780
ACGCGGCTTC CGCAGCTGAG GAAGTCCGAG AGCGAAGAGA GCAGGACCGC CAGGAGGCAG 840
ACCATCATCT TCCTGAGGCT GATTGTTGTG ACATTGGCCG TATGCTGGAT GCCCAACCAG 900
ATTCCGAGGA TCATGCTGTC GGCCAAACCC AAGCACGACT GGACGAGGTC CTACTTCCGG 960
GGGTACATGA TCCTCTCCCG CTTCCTGGAG ACGTTTTTCT ACCTCAGCTC GGTCTCAAC 1020
CCGCTCTCTG ACACGGTGTG CTGCGAGCAG TTTCGGGGGG TGTTCTGTGA GGTGCTGTGC 1080
TGCCGCTCTG CGCTGCAGCA CGCCCAACCAC GAGAAGCGCC TGCGCGTACA TGCGCACTCC 1140
ACCAACGACA GCGCCGCTTT TGTGCGAGCG CCGTTGCTCT TCGCGTCCCG GCGCCAGTCC 1200
TCTGCAAGGA GAACCTGAGAA GATTTTCTTA AGCACTTTTC AGAGCGAGGC CGAGCCCCAG 1260
TCTAAGTCCC AGTCATTGAG TCTCGAGTCA CTAGAGCCCA ACTCAGGGGC GAAACCAGCC 1320
AATTCGTCTG CAGAGAAATG TTTTCAGGAG CATGAAGTTT GA

Seq ID NO: 399 Protein sequence
Protein Accession #: NP_001499.1

1 11 21 31 41 51
MASPSLPGSD CSQIIDHSHV PEPFVATWIK ITLILVYLII FVMGLLGNV TIRVTQVLQK 60
KGYLQKEVTD HNVSLACSDI LVFLIGMPME FYSIIWNPLT TSSYTLCKL HTFLFEACSY 120
ATLLHVLTLF FERYIAICHP FRYKAVSGPC QVKLLIGFVW VTSALVALPL LFMGTVEYPL 180
VNVPSHRLGT CNRSSTRHHE QPETSNSMIC TNLSSRWTFV QSSIPGAFV VLVLLSVAP 240
MCWNMMQVLM KSKQKSLAGG TRPPQLRKSE SEESRTARRQ TIIIFLRLLV TLAVCWMPNQ 300
IRRMMAAKP KHDWTRSYFR AYMLLPFSE TFFYLSVVIN PLLYTVSSQ FRRVFVQVLC 360
CRLSLQHANH EKRLRVHAHS TTDSARFVQR PLLFASRRQS SARRTKIFL STFQSEAEFQ 420
SKSQSLSLSE LEPNSGAKPA NSAAENGPOE HEV

Seq ID NO: 400 DNA sequence
Nucleic Acid Accession #: NM_006475.1
Coding sequence: 28..2538

1 11 21 31 41 51
AACAGAACTG CAACGGAGAG ACTCAAGATG ATTCCCTTTT TAACCATGTT TCTCTACTA 60
TTGCTGCTTA TTGTTAACCC TATAAACGCC AACAAATCATT ATGACAAGAT CTGGCTCAT 120
AGTCGTATCA GGGGTGCGGA CCAAGGCCCA AATGTCTGTG CCCTTCAACA GATTTTGGGC 180
ACCAAAAAGA AATACCTTCAG CACTTGTAAAG AACTGGTATA AAAAGTCCAT CTGTGGACAG 240
AAAACGACTG TTTTATATGA ATGTTGCCCT GGTATATGA GAATGGAAGG AATGAAAGGC 300
TGCCCGAGCA TTTTGCCCAT TGACCATGTT TATGGCACTC TGGGCATCGT GGGAGCCACC 360
ACAACGCGAG GCTATTCTGA CGCCTCAAAA CTGAGGAGG AGATCGAGG AAAGGGATCC 420
TTCACCTACT TTGACCGAG TAATGAGGCT TGGGACAACT TGGATTCTGA TATCCGTAGA 480
GGTTTGGAGA GCAACGTGAA TGTGAAATTA CTGAATGCTT TACATAGTCA CATGATTAT 540
AAGAGAAATG TGACCAAGGA CTTAAAAAAT GGCATGATTA TTCCTTCAAT GTATAACAA 600
TTGGGGCTTT TCATTAACCA TTATCCTAAT GGGGTGTCTA CTGTTAAATG TGCTCGAATC 660
ATCCATGGGA ACCAGATTCA AACAAATGGT GTTGTCCATG TCATTGACCG TGTGCTTACA 720
CAAAATGGTA CCTCAATTCA AGACTTCATT GAAGCAGAAG ATGACCTTTC ATCTTTTAGA 780
GCAGCTGCCA TCACATCGGA CATATTGGAG GCCTTGGGAA GAGACGGTCA CTTCACACTC 840
TTTGCTCCCA CCAATGAGGC TTTTGAGAAA CTTCACGAG GTGTCTAGA AAGGTTTCATG 900
GGAGACAAAG TGGCTTCGA AGCTCTTATG AAGTACCACA TCTTAATAC TCTCCAGTGT 960
TCTGAGTCTA TTATGGGAGG AGCAGTCTTT GAGACGCTGG AAGGAAATAC AATTGAGATA 1020

5 GGATGTGACG GTGACAGTAT AACAGTAAAT GGAATCAAAA TGGTGAACAA AAAGGATATT 1080
 GTGACAAATA ATGGTGTGAT CCATTTGATT GATCAGGTCC TAATTCCTGA TTCTGCCAAA 1140
 CAAAGTTATTG AGCTGGCTGG AAAACAGCAA ACCACCTTCA CGGATCTTGT GGCCCAATTA 1200
 GGCTTGGCAT CTGCTCTGAG GCCAGATGGA GAATACACTT TGCTGGCACC TGTGAATAAT 1260
 GCATTTTCTG ATGATACTCT CAGCATGGTT CAGCGCTCC TTAATAAT TCTGCAGAA 1320
 CACATATTGA AAGTAAAGT TGGCCTTAAT GAGCTTTACA ACGGGCAAT ACTGGAAC 1380
 ATCGGAGGCA AACAGCTCAG AGTCTTGTGA TATCGTACAG CTGCTGCGAT TGAAAATTCA 1440
 TGCATGGAGA AAGGAGTAA GCAAGGGAGA AACGGTGC GA TACATATT CCGCGAGATC 1500
 10 ATCAAGGCAG CAGAGAATCT CCTCCATGAA AAGTTAAAC AAGATAAGCG CTTTAGCACC 1560
 TTCTCAGCC TACTTGAAGC TGCAGACTTG AAAGAGCTCC TGACACAACC TGGAGACTGG 1620
 ACATTATTG TGCCAAACCA TGATGCTTTT AAGGGAATGA CTAGTGAAGA AAAAGAAATT 1680
 CTGATACGGG ACAGAAATGC TCTTCAAAAC ATCATTCTTT ATCACCTGAC ACCAGGAGTT 1740
 TTCATTGGAA AAGGATTGGA ACCTGGTGT ACTAACATT TAAAGACCAC ACAAGGAAGC 1800
 15 AAAATCTTTC TGAAGAAGT AAATGATACA CTTCTGGTGA ATGAATTGAA ATCAAAAGAA 1860
 TCTGACATCA TGACAACAAA TGGTGTAAAT CATGTTGTAG ATAACTCTCT CTATCCAGCA 1920
 GACACACCTG TTGGAAATGA TCAACTGCTG GAAATACTTA ATAAATTAAT CAAATACATC 1980
 CAAATTAAGT TTGTTCTGAG TAGCACCTTC AAAGAAATCC CCGTGACTGT CTATACAAC 2040
 AAAATTATTA CCAAGTTGT GGAACCAAAA ATTAAGTGA TTGAAGGCAG TCTTCAGCCT 2100
 20 ATTATCAAAA CTGAAGGACC CACACTAACA AAAGTCAAAA TTGAAGTGA ACCTGAATTC 2160
 AGACTGATTA AAGAAGGTGA AACAATAACT GAAGTGATCC ATGGAGAGCC AATTATTAAA 2220
 AAATACACCA AAATCATTTGA TGGAGTGCT GTGGAATAA CTGAAAAGA GACACGAGAA 2280
 GAACGAATCA TTACAGGCTC TGAAATAAAA TACACTAGGA TTTCTACTGG AGGTGGAGAA 2340
 ACAGAAGAAA CTCTGAAGAA ATTGTTACAA GAAGAGGTCA CCAAGGTGAC CAAATTCATT 2400
 25 GAAGGTGGTG ATGGTCATT ATTGGAAGT GAAGAAATTA AAAGACTGCT TCAGGAGAC 2460
 ACACCGTGA GGAAGTTGCA AGCCAACAAA AAAGTTCAAG GTTCTAGAAG ACGATTAAGG 2520
 GAAGGTGCTT CTGAGTAAA ATCCAAAAAC CAGAAAAA TGTTTATACA ACCCTAAGTC 2580
 AATAACCTGA CCTTCAAGAA TTGTGAGAGC CAAGTTGACT TCAGGAAGCT AAACATCAGC 2640
 ACAGAGAGC AATCATCAAA TAATCTGAA CACAAATTTA ATATTTT TCTGTAATGA 2700
 30 GAAACATGAG GAAATTTG GAGTTAGCCT CCTGTGTTAA AGGAATTGAA GAAATATTA 2760
 CACCTTACAG CTTTTCAT CTTGACATTA AAAGTTCTGG CTAACCTTGG AATCCATTAG 2820
 AGAAAAATCC TTGTACACAG ATTCTATCA ATTCAATCG AAGAGTTGTG AACTGTTATC 2880
 CCATTGAAA GACCGAGCCT TGTATGTATG TTATGGATAC ATAAATGCA CGCAAGCAT 2940
 35 TATCTCTCCA TGGGAAGCTA AGTTATAAAA ATAGGTGCTT GGTGTACAAA ACTTTTATA 3000
 TCAAAAGGCT TTGCACATT CTATATGAGT GGGTTTACTG GTAAATTATG TTATTTTATA 3060
 CAACTAATT TGTACTTCA GAATGTTGT CATATGCTTC TTGCAATGCA TATTTTATA 3120
 TCTCAACGT TTCAATAAAA CCATTTTCA GATATAAAGA GAATTACTTC AAATTGAGTA 3180
 ATTCAGAAA ACTCAAGATT TAAGTAAAA AGTGGTTTGG ACTTGGGAA

Seq ID NO: 401 Protein sequence

Protein Accession #: NP_006466.1

40
 1 11 21 31 41 51
 45 MIPFLPMFSL LLLLVNPN ANHYDKILA HSRIRGRDQS PNVCALQOIL GTKKKYFSTC 60
 KNWYKKSICG QKTVLYECC PYMRMEGK GCPAVLPIDH VYGLGIVGA TTTQRYSDAS 120
 KLREEIEGKG SFYTFAPNE AWDNLDSDIR RGLSENVNVE LLNALHSHMI NKRMILTDLK 180
 NGMIIPSMYN NLGLFINHP NGVVTNVCAR IIEGNQIATN GVHVVIDRVL TQIGTSIQDF 240
 IEADDLSSF RAAAITSDIL EALGRDGHFT LPAPTNEAPE KLPRGVLERF MGDKVASEAL 300
 50 MKYHILNTIQ CSSEIMGAV FETLEGNTIE ICGDGDSTIV NGIKMVNKD IVTNNGVIHL 360
 IDQVLIPDSA KQVIELAGQV QTTTDLVAQ LGLASALRPD GEYTLAPVN NAFSDDTLSM 420
 VORLLKLIQ NHILKVKVL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKSKQ 480
 RNGAIHFRE IIKPAEKSLH EKLKQDKRPS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA 540
 PKGMTSEEKE ILIRDKNALQ NIIYLHLPF VPIGKGFEPG VINILKTTQG SKIFLKEVND 600
 55 TLLVNLKSK ESDIMTNGV IHVVDKLLYP ADTPVGNQDL LEILNKLIKY IQIKFVRGST 660
 FKBIPTVVT TKLIITVVEP KIKVIEGSLQ PIIKTEGPTL TKVKIEGEPE FRLIKEGETI 720
 TEVIEGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL 780
 QEEVTKVTKF IEAGDGHLFE DEEIKRLLQ DTPVRKLQAN KKVQGSRRRL REGRSQ

Seq ID NO: 402 DNA sequence

Nucleic Acid Accession #: NM_002416

Coding sequence: 40..417

60
 1 11 21 31 41 51
 65 ATCCAATACA GGAGTGACTT GGAACCTCAT TCTATCACTA TGAAGAAAAG TGGTGTCTT 60
 TTCTCTTGG GCATCATCTT GCTGGTCTG ATTGGAGTGC AAGGAACCCC AGTAGTGAGA 120
 AAGGGTCGCT GTTCCTGCAT CAGCACCAAC CAAGGAGTCA TCCACCTACA ATCCTTGAAA 180
 GACCTTAAAC AATTGCCCC AAGCCCTTCC TGCAGAAAAA TTGAAATCAT TGCTACACTG 240
 70 AAGAAATGGAG TTCAACATG TCTAAACCCA GATTTCAGAG ATGTGAAGGA ACTGATTAAA 300
 AAGTGGGAGA AACAGGTCAG CCAAAAGAAA AAGCAAAAGA ATGGGAAAAA ACATCAAAAA 360
 AAGAAAGTTC TGAAGTTG AAAATCTCAA CGTTCTCGTC AAAAGAAGAC TACATAAGAG 420
 ACCACTTAC CAATAAGTAT TCTGTGTTAA AAATGTTCTA TTTTAATTAT ACCGCTATCA 480
 TTCCAAAGGA GGATGGCATA TAATACAAAG GCTTATTAAT TTGACTAGAA AATTAAAAAC 540
 75 ATACTCTGA AATGTAACT AAAGTTAGAA AGTTGATTTT AAGAAATCAA ACGTTAAGAA 600
 TTGTTAAAG CTATGATTGT CTTTGTCTT CTACCAACCA CCAGTTGAAT TTCATCATGC 660
 TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT GTTCACCCAA CCACATCCCA 720
 CTCACACAG CTGCTGGAA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCAGAGAG 780
 TATCTGAGC ACATGTCAGC AAGTCCTAAG CTTGTTAGCA TGCTGGTGG CCAAGCAGTT 840
 80 TGAATTTGAG CTGACCTCA CCAAGCTGCT GTGGCCATCA ACCTCTGTAT TTGAATCAGC 900
 CTACAGGCCT CACACAAT GTGCTGAGA GATTCTATG GATTTGTTAT GGGTATCACC 960
 ACTGGAGATC ACCAGTGTGT GGCCTTCAGA GCCTCCCTTC TGGCTTTGGA AGCCATGTGA 1020
 TTCCATCTG CCGCTCAGC CTGACCACTT TATTTCTTTT TGTTCCTCTT TGCTTCATTC 1080
 AAGTCAGCTC TCTCCATCC TACCACAATG CAGTGCCCTT CTCTCTCCA GTGCACCTGT 1140
 85 CATATGCTCT GATTATCTG AGTCAACTCC TTTCTCATCT TGTCCCCAAC ACCCCACAGA 1200
 AGTGTCTTCT TCTCCAAAT CATCTCACT CAGTCCAGCT TAGTTCAAGT CCGCTCTCTT 1260
 AAATAAACCT TTTTGGACAC ACAAAATTAT TAAAACTCC TGTTCACCTT GGTTCAGTAC 1320
 CACATGGGTG AACACTCAAT GGTAACTAA TTCTTGGGTG TTTATCTCAT CTCTCCAAAC 1380

5 AGATTGTGAC CTCCTTGAGG GCAAGAGCCA CAGTATATTT CCCTGTTTCT TCCACAGTGC 1440
 CTAATAATAC TGTGGAACTA GGTTTTAATA ATTTTTTAAT TGATGTTGTT ATGGGCAGGA 1500
 TGGCAACACG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCCTGGCTAC TCCATGTTGG 1560
 CTAGCCTCTG GTAACTCTT ACTTATTATC TTCAGGACAC TCACTACAGG GACCAGGGAT 1620
 GATGCAACAT CCTTGTCTTT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA 1680
 GCACGTGGTA AAACACTTGC GGATATTCTG GACTGTTTTT AAAAAATATA CAGTTTACCG 1740
 AAAAAATATA AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC 1800
 CCAACCATAC AAAAAATCTT TTTCCGGAAG GAAAAGGGCT TTCTCAATAA GCCTCAGCTT 1860
 TCTAAGATCT AACAAAGATG CCAACGAGAT CCTTATCGAA ACTCATTTTA GGCAAAATAG 1920
 10 AGTTTTATTG TCCGTTTACT TGTTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA 1980
 TCTCCCATGA AGAAAGGGAA CGGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT 2040
 TAGTGAAGC ATGATTGGTG CCCAGTTAGC CTCTGCAGGA TGTGGAAACC TCCTTCCAGG 2100
 GGAGGTTTCA TGAATTGTGT AGGAGAGGTT GTCTGTGGCC AGAATTTAAA CCTATACCTA 2160
 CTTCCTCAAA TTGAATCACT GCTCACACTG CTGATGATT AGAGTGCTGT CCGGTGGAGA 2220
 15 TCCACCCGGA AGCTCTTATC TAATCATGAA ACTCCCTAGT TCCTTCATGT AACTTCCCTG 2280
 AAAAAATCTA GTGTTTCATA AATTGAGAG TCTGTGACCC ACTTACCTTG CATCTCACAG 2340
 GTAGACAGAT TATAACTTAA AACCAAGAC TACATATTGT CACTGACACA CACGTTATAA 2400
 TCAATTTATCA TATATATACA TACATGCATA CACTCTCAA GCAAAATAAT TTTCACTTCA 2460
 20 AAACAGTATT GACTTGTATA CCTTGTAAAT TGAATATATT TCTTTGTTAA AATAGAAATG 2520
 TATCAATAAA TAGACCATTA ATCAG

Seq ID NO: 403 Protein sequence
 Protein Accession #: NP_002407

25 1 11 21 31 41 51
 MKKSGVLFLD GIILLVLIGV OGTFVVRKGR CSCISTNQGT IHLQSLKDLK QFAPSPSCEK 60
 IEIIATLKNG VQTCLNPDSA DVKELIKKWE KQVSQKKKQK NGKKHQKKKV LKVRKSQRSR 120
 QKKT

Seq ID NO: 404 DNA sequence
 Nucleic Acid Accession #: NM_006670
 Coding sequence: 85..1347

35 1 11 21 31 41 51
 CCGGCTCGGG CCTTCGGGCG CAGCCTTCCC GAGCCTTCGG AGCGGGGCGC GTCCAGCCCC 60
 AGCTCCGGGG AAACGCGAGC CGCGATGCTT GGGGGGTGCT CCGGGGGCCC CGCGCGCGGG 120
 40 GAGCGGGGTC TGGCGGTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCTCTCGTCT 180
 TCTCCCACTC CCTCGGCATC CTCTCTCTCC TCCTCGGCGC CGTTCTCTGG TTCCGCGCGT 240
 TCGCGCCAGC CCGCGCTGCC GGACCAAGTG CCGCGCTGT GCGAGTGCTC CGAGGCAGCG 300
 CGCACAGTCA AGTGCCTTAA CCGCAATCTG ACCGAGGTGC CCACGGAACCT GCCCGCTAC 360
 GTGCGCAACC TCTTCTTAC CCGCAACCAG CTGGCCGTGC TCCCTGCGCG CGCCTTGGCC 420
 CGCGCGCGCG CGCTGGCGGA GCTGGCGCGC CTCAACCTCA GCGCGAGCCG CCTGGAAGAG 480
 45 GTGCGCGCGG CGCCTTTCGA GCATCTGCCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC 540
 CCACTGGCGC ACCTCAGTCC CTTCGCTTTC TCGGGCAGCA ATGCCAGCGT CTCGGCCCCC 600
 AGTCCCTCTG TGGAACTGAT CCTGAACCA ATCGTGCCCC CTGAAGATGA GCGGCAGAAC 660
 CGGAGCTTGG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCCGTGCACT GCAGGGGCTC 720
 CGCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCAA 780
 50 CTGCGCCAGC TCAGGCACCT GGACTTAAAT AATAATTGCG TGGTGAGCCT GACCTACGTG 840
 TCCTTCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900
 CTTCAATAAG GCACCTCTGG TGAGTTGCAA GGTCTACCCC ACATTAGGTT TTTCTGGAC 960
 AACAAATCCCT GGGTCTGCGA CTGCCACATG GCAGACATGG TGACCTGGCT CAAGGAAACA 1020
 GAGGTAGTGC AGGGCAAAAG CCGGCTCACC TGTGCATATC CGGAAAAAAT GAGGAATCGG 1080
 55 GTGCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCACTCCCTG 1140
 CAAACCTCTT ATGTCTTCTC GGGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCTCCTG 1200
 GTTTTGTATT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCGCTG 1260
 AGGGATCACA TGGGAAGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAAAC 1320
 60 AACCTCAGTT CTAACCTCGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACCTCT 1380
 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA 1440
 TAGATACAAC GGACTTTGAC TAAAGCAGT GAAGGGGATT TGCITCTCTG TTATGTAAG 1500
 TTTCTCGGTG TGTCTGTGTA ATGTAAGAG ATGAACAGTT GTGTATAGTG TTTTACCTC 1560
 TTTCTTTTCT TGGAACTCCT CAACACGTAT GGAGGGATT TTCAGGTTTC AGCATGAACA 1620
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680
 65 ACAGATAGCA TTCAACAAAA GCTGCTCTCA CTTTTTCGAG AAAAAATCTT TATTCAATAA 1740
 TATCAGTTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAT AATTGCATCC TATAAACTGC 1800
 CTGCAGACGT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA 1860
 AGAGCATGCT TACATTTTAC TGTCTGCTAT ATTACAAAAA ATAACCTTGA ACTTCATAAC 1920
 70 TTTCTTGACA AAGTAAATTA CTTTTTGAT TGCAGTTTAT ATGAAAATGT ACTGATTTTT 1980
 TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAAATAAG 2040
 ATTCTTAAAA GAA

Seq ID NO: 405 Protein sequence
 Protein Accession #: NP_006661

75 1 11 21 31 41 51
 MPGGCSRGA AGDRLRLAR LALVLLGWVS SSSPTSSASS FSSAPFLAS AVSAQPPLPD 60
 80 QCPALCBCEG AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL 120
 AALNLSGSLR DEVRAGAFEH LPSLRQLDLS HNPLADLSPF AFGSNAASVS APSPLVELIL 180
 NHIVPPEDER QNRSFEGMVV AALLAGRALQ GLRRLELASN HFLVLPRLVL AQLPSLRHLD 240
 LSNLSLVSLT YVSRNLTHL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVP LDNNPFWCDC 300
 HMDMVTWLK ETEVVGKDR LTCAYPEKMR NRVLLELNSA DLDGDPILPP SLQTSYVFLG 360
 85 IVLALIGAIF LLVLYLNRKG IKKWMHNIRD ACRDHMBGYH YRYEINADPR LTNLSSNSDV

Seq ID NO: 406 DNA sequence
 Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..927

	1	11	21	31	41	51	
5	ATGCTCTGGG	GCTGCTCCCG	GGGCCCCCGCC	GCCGGGGAGC	GGCGTCTGCG	GCTGGCGCGA	60
	CTAGCGCTGG	TACTCTCTGG	CTGGGTCTCC	TCGTCTTCTC	CCACCTCCTC	GGCATCTCTCC	120
	TTCTCTCTCT	CGCGCGCGTT	CCTGGCTTCC	GCCGTGTCCG	CCCAGCCCCC	GCTGCCGGAC	180
	CAGTGCCTCC	CGCTGTGGGA	GTGCTCCGAG	GCAGCGCGCA	CAGTCAAGTG	CGTTAACCAG	240
	AATCTGACCG	AGGTGCCAC	GGACCTGCCC	GCCTACGTGC	GCAACCTCTT	CCTTACCGGC	300
10	AACCAGCTGG	CCAGCAACCA	CTTCTTTTAC	CTGCCGCGGG	ATGTGCTGGC	CCAACCTGCC	360
	AGCCTCAGCG	ACCTGGACTT	AAGTAATAAT	TCGTCTGTGA	GCCTGACCTA	CGTGTCTTTC	420
	CGCAACCTGA	CACATCTAGA	AAGCCTCCAC	CTGGAGGACA	ATGCCCTCAA	GGTCTCTCAC	480
	AATGGCACCC	TGGCTGAGTT	GCAAGGTCTA	CCCCACATTA	GGGTTTTCCT	GGACAACAAT	540
	CCCTGGGTCT	GCGACTGCCA	CATGGCAGAC	ATGGTGACCT	GGCTCAAGGA	AACAGAGGTA	600
15	GTGCAGGGCA	AAGACCGGCT	CACCTGTGCA	TATCCGGAAA	AAATGAGGAA	TCGGGTCTCT	660
	TTGGAAGTCA	ACAGTGTGTA	CCTGGACTGT	GACCCGATTC	TTCCCCATC	CCTGCAAAAC	720
	TCTTATGTCT	TCTGTGGTAT	TGTTTTAGCC	CTGATAGGCG	CTATTTTCCT	CCTGGTTTGT	780
	TATTTGAACC	GCAAGGGGAT	AAAAAAGTGG	ATGCATAACA	TCAGAGATGC	CTGCAGGGAT	840
20	CACATGGAAG	GGTATCATTA	CAGATATGAA	ATCAATGCGG	ACCCAGATT	AACAAACCTC	900
	AGTTCTAACT	CGGATGTCTT	CGAGTGA				

Seq ID NO: 407 Protein sequence
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
25	MPGGCSRPGA	AGDGRLLRLR	LALVLLGWVS	SSSPSTSSASS	FSSSAPFLAS	AVSAQPPLPD	60
	QCPCALCESE	AARTVKCVNR	NLTVPTDLF	AYVRNLFITG	NQLASNHFLY	LEPDVLAQLP	120
	SLRHLDSLNN	SLVSLTYVSP	RNLTHLESLH	LEDNALKVLH	NGTLAEQLGL	PHIRVFLDNN	180
30	PWVCDCHMAD	MVTWLKETEY	VQKDRILTCA	YPERKMRNRVL	LELNSADLDC	DPILPPSLQT	240
	SYVFLGIVIA	LIGAIFFLLV	YLNKGIKKW	MHNIRDACRD	HMEGYHYRYE	INADPRLTNL	300
	SSNSDVLE						

Seq ID NO: 408 DNA sequence
Nucleic Acid Accession #: NM_000095.1
Coding sequence: 26..2299

	1	11	21	31	41	51	
40	CAGCACCCAG	CTCCCGGCCA	CCGCCATGGT	CCCCGACACC	GOCTGCGTTC	TCTGCTCAC	60
	CCTGGCTGCC	CTCGGCGCGT	CCGGACAGGG	CCAGAGCCCG	TTGGGCTCAG	ACCTGGGCCC	120
	GCAGATGCTT	CGGGAACCTG	AGGAAACCAA	CGCGGCGCTG	CAGGAAGTGC	GGGACTGGCT	180
	GCGGACGACG	GTCAAGGAGA	TCACGTTTCT	GAAAAACACG	GTGATGGAGT	GTGACGCGTG	240
	CGGGATGACG	CAGTCAGTAC	GCACCGGCCT	ACCCAGCGTG	CGGCCCTGTC	TCCACTGCGC	300
45	GCCCGGCTTC	TGCTTCCCCG	CGGTGGCCTG	CATCCAGACG	GAGAGCGGGC	GCCGCTGCGG	360
	CCCTGCGCCC	GCGGGCTTCA	CGGGCAACGG	CTCGCACTGC	ACCCAGCTCA	ACGAGTGCAA	420
	CGCCACACCC	TGCTTCCCCG	GAGTCCGCTG	TATCAACACC	AGCCCGGGGT	TCCGCTGCGA	480
	GGCTTGCCCG	CCGGGGTACA	GCGGCCCCAC	CCACCAGGGC	GTGGGGCTGG	CTTTGCGCAA	540
	GGCCAAACAG	CAGGTTTGCA	CGGACATCAA	CGAGTGTGAG	ACCGGGCAAC	ATAACTGCGT	600
50	CCCCAACTCC	GTGTGCATCA	ACACCCGGGG	CTCCTTCCAG	TGCGGCGCGT	GCCAGCCCGG	660
	CTTCGTGGGC	GACCAAGCGT	CCGGCTGCCA	GCGCGGCGCA	CAGCGCTTCT	GCCCGGACGG	720
	CTGCGCCACG	GAGTGCACAG	AGCATGCAGA	CTGCGTCTTA	GAGCGCGATG	GCTCGCGGTC	780
	GTGCGTGTGT	CGCGTGGCTT	GGGCCGGCAA	CGGGATCTCT	TGTGGTGGCG	ACACTGACCT	840
	AGACGGCTTC	CCGGAGCAGA	AGCTGCGCTG	CCCGAGCCCG	CAGTGCCTGA	AGGCAAACTG	900
55	CGTGACTGTG	CCCAACTCAG	GGCAGGAGGA	TGTGGACCGC	GATGGCATOG	GAGACGCTGT	960
	CGATCCGGAT	GCCGACGGGG	ACGGGGTCCC	CAATGAAAAG	GACAACTGCC	CGCTGGTGGG	1020
	GAACCCAGAC	CAGCGCAACA	CGGACGAGGA	CAAGTGGGGC	GATGCGTGGG	ACAACTGGCG	1080
	GTCCCAAGAG	AACGACGACC	AAAAGGACAC	AGACCAGGAC	GGCCGGGGCG	ATGCGTGGCA	1140
	CGACGACATC	GAGCGGCGAC	GGATCCGCAA	CCAGGCGGAC	AACTGCCCTA	GGGTACCCAA	1200
60	CTCAGACGAG	AAGGACAGTG	ATGGCGATGG	TATAGGGGAT	GOCTGTGACA	ACTGTCCCCA	1260
	GAAGAGCAAC	CCGGATCAGG	CGGATGTGGA	CCACGACTTT	GTGGGAGATG	CTTGTGACAG	1320
	CGATCAAGAC	CAGGATGGAG	ACGGACATCA	GGACTCTCGG	GACAACTGTC	CCACGGTGCC	1380
	TAACAGTGCC	CAGGAGGACT	CAGACCACGA	TGGCCAGGGT	GATGCTTGGG	ACGACGACGA	1440
	CGACATGAC	GGAGTCCCTG	ACAGTCCGGA	CAACTGCCCG	CTGGTGCTTA	ACCCCGGCCA	1500
65	GGAGGACCGG	GACAGGACCG	GCGTGGGCGA	CGTGTGCCAG	GACGACTTTG	ATGCAGACAA	1560
	GGTGTAGAC	AAGATCGACG	TGTGTCCGGA	GAACGCTGAA	GTCAAGCTCA	CCGACTTCAG	1620
	GGCCTTCCAG	ACAGTGTGTC	TGGACCCGGA	GGGTGACGCG	CAGATTGACC	CCAACTGGGT	1680
	GGTGTCTCAAC	CAGGGAAGGG	AGATCGTGCA	GACAAATGAA	AGCGACCCAG	GCCTGGCTGT	1740
	GGGTACTACT	GCCTTCAATG	GCGTGGACTT	CGAGGGCAGG	TTCCATGTGA	ACACGCTCAC	1800
70	GGATGACGAC	TATGCGGGCT	TCATCTTTGG	CTACCAGGAC	AGCTCCAGCT	TCTACGTGGT	1860
	CATGTGGAAG	CAGATGGAGC	AAACGTATTG	GCAGGGGAAAC	CCCTTCCGTG	CTGTGGCCGA	1920
	GCCTGGCATC	CAACTCAAGG	CTGTGAAGTC	TTCCACAGGC	CCCGGGGAAC	AGCTGCGGAA	1980
	CGCTCTGTGG	CATACAGGAG	ACACAGAGTC	CCAGGTGCGG	CTGCTGTGGA	AGGACCGCGG	2040
	AAACGTGGGT	TGGAAGGACA	AGAAGTCCTA	TCGTGTGGTC	CTGCAGCACC	GGCCCCAAGT	2100
75	GGGCTACATC	AGGGTGGGAT	TCTATGAGGG	CCCTGAGCTG	GTGGCGGACA	GCAACGTGGT	2160
	CTTGACACAC	ACCATGCGGG	GTGGCGGCTT	GGGGGTCTCT	TGCTTCTCCC	AGGAGAACAT	2220
	CATCTGGGCC	AACCTGCGTT	ACCGCTGCAA	TGACACCATC	CCAGAGGACT	ATGAGACCCA	2280
	TCAGCTGGCG	CAAGCTTAGG	GACCAGGGTG	AGGACCCGCG	GGATGACAGC	CACCTCACC	2340
80	GCGGCTGGAT	GGGGGCTCTG	CACCCAGCCC	AAGGGGTGGC	CGTCTGAGG	GGGAAGTGAG	2400
	AAGGGCTCAG	AGAGACAAA	ATAAAGTGTG	TGTGCAGGG			

Seq ID NO: 409 Protein sequence
Protein Accession #: NP_000086.1

	1	11	21	31	41	51	
85	MVPDTACVLL	LTLAALGASG	QGQSPLGSDL	GPQMLRELQE	TNAALQDVDR	WLRQQVREIT	60

FLKNTVMED ACQMQQSVRT GLPSVRPLLE CAPGFCFPGV ACIQTESGGR CGPCPAGFTG 120
NGSHCTDVNE CNAHPCFPRV RCINTSPGFR CEACPPGYSG PTHQGVGLAF AKANKQVCTD 180
INECETGOHN CVPNSVCINT RGSFQCGPCQ PGFVGDQASG CQRGAQRPCP DGSPSECHEH 240
ADCVLERDGS RSCVCRVGWA GNGILCGRDT DLDGFPDEKL RCPEPQCRKD NCVTVPNSGQ 300
EDVDRDGDG ACDFDADGDG VPNEKDNCPV VRNPDQRTND EDKWDGACDN CRSQKNDQDK 360
DTQDQGRGDA CDDIDIDGRI RNQADNCPRV FNSDQKSDG DGIGDADCMC PQKSNPDQAD 420
VDHDFVGDAC DSDQDQDGDG HQDERDNCPV VFNQAQEDSD HDGQGDACDD DDDNDGVFDS 480
RDNCRLVNPV GQEDADRDRGVD GDVCQDDFDA DKVVDKIDVC PENAEVLTLD FRAFQTVVLD 540
PEGDAQIDFN WVVLNQGRIE VQTMNSDPGL AVGYTAFNGV DFEGTFHVTI VTDDDYAGFI 600
FGYQDSSSFY VVMWQMEQET YWQANPFRV AEPGIQLKAV KSSTGPGQEL RNALWHTGDT 660
ESQVRLMKD PRNVGKDKK SYRNFLQHRP QVGIIRVRFY EGPELVADSN VVLDTTMRGG 720
RLGVPCFSQE NIHWANLRYR CNDTIPEDYE THQLRQA

Seq ID NO: 410 DNA sequence
Nucleic Acid Accession #: NM_001565.1
Coding sequence: 67..363

1 11 21 31 41 51
GAGACATTC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC CTCCAGTCTC 60
AGCACCATGA ATCAACTGCG GAITCTGATT TGCTGCTCTA TCTTTCTGAC TCTAAGTGGC 120
ATTCAAGGAG TACCTCTCTC TAGAACCGTA CGCTGTACCT GCATCAGCAT TAGTAATCAA 180
CCTGTTAATC CAAGGTCTTT AGAAAACTT GAAATTATTC CTGCAAGCCA ATTTTGTCCA 240
CGTGTGAGA TCATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA 300
TCGAAGGCCA TCAAGAAATT ACTGAAAGCA GTTAGCAAGG AAATGTCTAA AAGATCTCCT 360
TAAACCCAGA GGGGAGCAAA ATCGATGCGA TGCTTCCAAG GATGGACCAC ACAGAGGCTG 420
CCTCTCCCAT CACTTCCCTA CATGGAGTAT ATGTCAAGCC ATAATTGTTC TTAGTTTGCA 480
GTACACTAA AAGGTGACCA ATGATGGTCA CCAATCAGC TGCTACTACT CCTGTAGGAA 540
GGTTAATGTT CATCATCTTA AGCTATTCTA TAATAACTCT ACCCTGGCAC TATAATGTAA 600
GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTTCTGACC CTGCTTCAA TATTTCCTCT 660
ACCTTTCCCA TCTTCCAAGG GTACTAAGGA ATCTTTCTGC TTTGGGGTTC ATCAGAATTC 720
TCAGAATCTC AAATAACTAA AAGGTATGCA ATCAATCTG CTTTTTAAAG AATGCTCTTT 780
ACTTCATGGA CTTCACCTGC CATCTCCCA AGGGGCCCAA ATTCCTTCAG TGGCTACCTA 840
CATACAAATC CAACACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAAGTATT 900
CTTATTTAAT GAAAGACTGT ACAAGATATA AGTCTTAGAT GTATATATTT CCTATTATGT 960
TTTCAGTGTG CATGGAATAA CATGTAATTA AGTACTATGT ATCAATGAGT AACAGGAAAA 1020
TTTTAAATAT ACAGATAGAT ATATGCTCTG CATGTTACAT AAGATAAATG TGCTGAATGG 1080
TTTTCAATAA AAAATGAGGT ACTCTCCTGG AAATATTAAG

Seq ID NO: 411 Protein sequence
Protein Accession #: NP_001556.1

1 11 21 31 41 51
MNQTALICC LIFLTLISGIQ GVPLSRTVRC TCISISNPQV NPSLEKLEI IPASQPCPRV 60
EIIATMKKKK EKRLNLPESK AIKNLLKAVS KEMSKRSP

Seq ID NO: 412 DNA sequence
Nucleic Acid Accession #: XM_057014
Coding sequence: 143..874

1 11 21 31 41 51
GGGAGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCTCGGAG 60
CGGGGCGGAG CCAGAGCGTG ACCACGTTCC TCTCCTCGGT CTCCTCGGCC TCCAGCTCCG 120
CGCTGCGCGC CAGCGCGGAG CCATGCGACC CCAGGGCCCC GCGGCTCTCC CGCAGCGGCT 180
CGCGGCGCTC AGCTGCTCTC TGCTGCTGCA GCTGCGCGCG CCGTCGAGCG CCTCTGAGAT 240
CCCCAAGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG 300
AATGTGCTTA CAAGGCGCCAG CAGGAGTGCC TGGTCGAGAC GGGAGCCCCG GGGCCAATGG 360
CATTGCGGGT ACACCTGGGA TCCAGGTGCG GGATGGATTC AAAGGAGAAA AGGGGGAATG 420
TCTGAGGGAA AGCTTTGAGG AGTCTGGAC ACCCAACTAC AAGCAGTGT CATGGAGTTC 480
ATTGAATTAT GGCATAGATC TTGGGAAAT TGCGGAGTGT ACATTATCAA AGATGCGTTC 540
AAATAGTGCT CTAAGAGTIT TGTTCACTGG CTCACCTCGG CTAATATGCA GAAATGCATG 600
CTGTGAGCGT TGGTATTTC AATCAATG AGCTGAATGT TCAGGACCTC TCCCATTTGA 660
AGCTATAATT TATTGGACC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTATCTG 720
CACTTCTTCT GTGGAAGGAG TTTGTGAAG AATTGGTGCT GGATTAGTGG ATGTTGCTAT 780
CTGGGTGGC ACTTGTTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT 840
TTCTCGCATC ATTATTGAAG AACTACCCAA ATAAATGCTT TAATTTTCAT TTGTACCTCT 900
TTTTTTTATT ATGCTTGGGA ATGTTTCACT TAAATGACAT TTTAAATAAG TTTATGTATA 960
CATCTGAATG AAAAGCAAAG CTAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT 1020
TTTAAATCTA GCATTATTCA TTTTGCTTCA ATCAAAAGTG GTTTCAATAT TTTTITTAGT 1080
TGGTTAGAAT ACTTCTTCCA TAGTCACATT CTCTCAACCT ATAATTGGA ATATTGTGT 1140
GGTCTTTTGT TTTTCTCTCT AGTATAGCAT TTTTAAAAA ATATAAAAGC TACCAATCTT 1200
TGTACAATTT GTAAATGTTA AGAATTTTTT TTATATCTGT TAAATAAAAA TTATTTCCAA 1260
CAACCTTAAA AAAAAAAAAA AAAA

Seq ID NO: 413 Protein sequence
Protein Accession #: XP_057014

1 11 21 31 41 51
MRPQGAASP ORLRGLLLLL LLQLPAPSSA SEIPKQKQKA QLRQREVVDL YNGMCLQGPA 60
GVPRDGSFPG ANGIPGTGPI PGRDGPKEK GECLRESFEE SWTPNYKQCS WSSLNVDGL 120
GKIAECTFTK MRNSALRVL FSGSLRLKCR NACQQRWYPT PNGAECSGPL PIEAIIVLDQ 180
GSPENNSTIN IHRSTSVEGL CEGIGAGLVD VAIWVGTCSD YPKGDASTGW NSVSRIIEE 240
LPR

WO 02/086443
Seq ID NO: 414 DNA sequence
Nucleic Acid Accession #: XM_084007
Coding sequence: 138..2405

PCT/US02/12476

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
CTGTGTCGGA ATTGGGCAGG AGACCGCGTG TTCGCGCTGT GTAGAGATTT CTCGAAGACA 60
CCATGGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCGG GGCCTGGGGA CAACGAGGCC 120
GCGGAGACGA AGCGGCAATG GCGAGGAAGT TATCTGTAAT CTGTATCTTG ACCTTTGGCC 180
TCTCTGTGAC AAATCCCCCT CATGAACATA AAGCAGCTGC TTTCCCCCAG ACCACTGAGA 240
AAATTAGTCC GAATTGGGAA TCTGGCATTG ATGTTGACTT GGCAATTTCC ACACGGCAAT 300
ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCCTTGTCA GTTGAAGGGT 360
TCAGAAAATT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG 420
ACCAAGACCA TCACTCAGAC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG 480
AGCATCACTC AGACCAAGAG CATCACTCTG ACCATGATCA TCACTCCAC CATAATCATG 540
CTGCTTCTGG TAAAAATAAG GAAAAAGCTC TTTGCCGAGA CCATGACTCA GATAGTTGAG 600
GTAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA 660
GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCACCTGTG TACAACACTG 720
TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACCTCTTC 780
CCAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCACTAT AAAGAGCCGG GTGAGCCGGC 840
TGGCTGTGAG GAAAAATAAT GAATCTGTGA GTGAGCCCGG AAAAGGCTTT ATGTATTCCA 900
GAAACACAAA TGAAATCCTT CAGGAGTGTT TCAATGCATC AAAGCTACTG ACATCTCATG 960
GCATGGGCAAT CAGGTTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA 1020
TCAACCAAAAT TGATGCTAGA TCTGTCTGTA TTCAACAAG TGAAAAAGAG GCTGAAATCC 1080
CTCAAAAGAC CTATTCATTA CAAATAGCCT GGGTTGTGGG TTTTATAGCC ATTTCCATCA 1140
TCAGTTTCTT GTCTCTGCTG GGGTTATCT TAGTGCCTCT CATGAATCGG GTGTTTTTCA 1200
AATTTCTCCT GAGTTTCTCT GTGGCACTGG CGGTGGGAC TTTGAGTGGT GATGCTTTTT 1260
TACACCTTCT TCACATTTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG 1320
CAATGGAAT TCAGTGAAGG CCACTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA 1380
GTGCCTATTT TGATCCACG TGGAAGGGTC TAACAGCTCT AGGAGGCCGT TATTTCATGT 1440
TTCTTGTTGA ACATGTCTCT ACATTGATCA AACAAATTAA AGATAAGAAG AAAAAGAATC 1500
AGAAGAATCC TGAAATGATG GATGATGTGG AGATTAGAA GCAGTTGTCC AAGTATGAAT 1560
CTCACTTTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAAGCTGA GGCTATTTC 1620
GAGCAGACTC ACAAGAGGCC TCCCACTTTG ATTCTCAGCA GCGTCAGCTC TTGGAAGAAG 1680
AAGAGGTGAT GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG 1740
GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA 1800
TTCAACACCA TCATGACTAC CATCATATT TCCATCATCA CCACCACCAA AACCACCATC 1860
CTCAGACTCA CAGCCAGCGC TACTCTCGGG AGGAGCTGAA AGATGCCCGC GTCGCCACTT 1920
TGGCCTGGAT GGTGATAAGT GGTGATGGCC TGCACAAATT CAGGATGGCC CTAGCAATTG 1980
GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTTGCT GTGTTCTGTC 2040
ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA 2100
AGCAGCTGTG CCTTTATAAT GCATTGTGAG CCATGCTGGC GTATCTTGA ATGGCAACAG 2160
GAATTTTCAT TGCTCATATT GCTGAAATG TTTCTATGTG GATATTGCA CTTACTGCTG 2220
GCTTATTATG GTATGTTGCT CTGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA 2280
GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTACA GAATGCTGGG ATGCTTTTGG 2340
GTTTGGAAAT TAGTTACTTT ATTTCCATAT TTGAACATAA AATCGTGTTC CGTATAAAT 2400
TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAAGTTG TCATAGTTTC AGTAGGTCAT 2460
AGGAGATGTA GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGGT TTTGTGATT 2520
TTGTATTGAA TATTGCTGTC TGTTACAAAG TCAGTTAAAG GTACGTTTAA ATATTTAAGT 2580
TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTACCC GGTATTACCA GTTTATTATG 2640
TAAACAGAGG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAATAATGT CTTTAAATGT 2700
TTTTCAAGAA CTAAACAGT TATTCCTATA CTGGATTATA GGTCTCTGAA GAACTGCTGG 2760
TGTTTAGGAA TAAGAAATGT CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA 2820
AGCAAGAGAA TAAAGAGAGG AAGAGAGAGG TCTGAGAAAT GGGGAGGCAT AGATTCTTAT 2880
AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAG TATAAAAGG 2940
CAGAATTAGT ATAGAGTACA TTCAATTAAC ATTTTGTGCA GGATTATTTC CCGTAAAAAC 3000
GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTAAAC TTTGTATAAT ACAGAAATCT 3060
AAATATATTT AATGAATTTA AGCAATATAC ACTTGACCAA GAAATTTGGA TTTCAAAATG 3120
TTGCTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT 3180
TATTGCCAAG TTATATATCA CCAAAGCTG TATGACTGGA TGTCTGTTT ACCTGGTTTA 3240
CAAAATTTAT AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAACT ACACTAAGTA 3300
TCATTTGATT CGATTAGAAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTTG 3360
GAGCAATTTT CTTTATATAC GGTACTGTAG CCATCTAGG CCTGCTGTG GCATCTCTTA 3420
GATGTTTCTT TTTTACAAA TAAATTCCTT ATATCAGCTT G

```

Seq ID NO: 415 Protein sequence
Protein Accession #: XP_084007

```

70      1      11      21      31      41      51
      |      |      |      |      |      |
MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPNN ESGINVDLAI STRQYHLQQL 60
FYRYGENNSL SVEGFRKLLQ NIGIDKIKRI HIHHDHHDHS DHEHSDHER HSDHEHSDH 120
EHSDDHDS HNHAASGN KRKALCPDHD SDSSGKDPNN SQKGARHPE HASGRNRVKD 180
SVSASEVTST VYNTVSEGTH PLETIETPRP GKLPKDVSS STPPSVTSKS RVSRLAGRKT 240
NESVSEPRKG FMYSRNTNEN PQECFNASKL LTHSGMGIQV PLNATEPNYL CPAIINQIDA 300
RSCLHTSEK KAEIPPKTYS LQIAWVGFI AISIISFLSL LGVILVPLMN RVFFKLLSP 360
LVALAVGTLG GDAFLHLLEH SHASHHSHS HEEFAMEMKR GPLFSLSSQ NIEESAYFDS 420
TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKNQKKPEN DDDVEIKKQL SKYESQLSTN 480
EEKVDTDDRT EGYLRADSQE PSHFDSQPPA VLEEEVMIA HAHPQEVYNE YVPRGCMKC 540
HSHFDITLQ SDLIHLHHDH YHILHHDH QNHHPHSHSQ RYSREELKDA GVATLAWMVI 600
MGDGLHNSD GLAIGAAFTG GLSSGLSTSV AVFCHPELPE LGDFAVLLKA GMTVKQAVLY 660
NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS 720
RWGYPLQNA GMLLPGIML LISIFERKIV FRINP

```

Seq ID NO: 416 DNA sequence
Nucleic Acid Accession #: NM_015419.1
Coding sequence: 1..8487

	1	11	21	31	41	51	
5	ATGCCCAAGC	GCGCGCACTG	GGGGGCGCTC	TCCGTGGTGC	TGATCCTGCT	TTGGGGCCAT	60
	CCGCGAGTGG	CGCTGGCCTG	CCGCGATCCT	TGTGCTCTGT	ACGTCCCCAG	CGAGGTCCAC	120
	TGCACGTTCC	GATCCCTGGC	TTCCGTGCCC	GCTGGCATTG	CTAGACACGT	GGAAAGAATC	180
	AATTTGGGGT	TTAATAGCAT	ACAGGCGCTG	TCAGAAACCT	CATTTCGAGG	ACTGACCAAG	240
	TTGGAGCTAC	TTATGATTCA	CGGCAATGAG	ATCCCAAGCA	TCCCGATGGG	AGCTTTAAGA	300
10	GACCTCAGCT	CTCTTCAGGT	TTTCAAGTTC	AGCTACAACA	AGCTGAGAGT	GATCACAGGA	360
	CAGACCTCC	AGGGTCTCTC	TAACTTAATG	AGGCTGCACA	TTGACCACAA	CAAGATCGAG	420
	TTTATCCACC	CTCAAGCTTT	CAACGGCTTA	ACGTCTCTGA	GGCTACTCCA	TTTGGAAAGG	480
	AATCTCCTCC	ACCAGCTGCA	CCCCAGCACC	TTCTCCACGT	TCACATTTT	GGATTATTTT	540
	AGACTCTCCA	CCATAAGGCA	CCTCTACTTA	GCAGAGAAAC	TGGTTAGAAC	TCTTCCTGCC	600
15	AGCATGCTTC	GGAACATGCC	GCTTCTGGAG	AATCTTTACT	TGCAGGGAAA	TCCGTGGACC	660
	TGCGATTGTG	AGATGAGATG	GTTTTTGGAA	TGGGATGCAA	AATCCAGAGG	AATCTGGAAG	720
	TGTAAAAAGG	ACAAAGCTTA	TGAAGGCGGT	CAGTTGTGTG	CAATGTGCTT	CAGTCCAAAG	780
	AAGTTGTACA	AACATGAGAT	ACACAAGCTG	AAGGACATGA	CTTGTCTGAA	GCCTTCAATA	840
	GAGTCCCTCC	TGAGACAGAA	CAGGAGCAGG	AGTATTGAGG	AGGAGCAAGA	ACAGGAAGAG	900
20	GATGGTGCCA	GCCAGCTCAT	CCTGGAGAAA	TTCCAACCTG	CCCAGTGGAG	CATCTCTTTG	960
	AATATGACCG	ACGAGACCGG	GAACATGGTG	AACCTTGGCT	GTGACATCAA	GAAACCAATG	1020
	GATGTGTACA	AGATTCACTT	GAACCAAAAG	GATCCTCCAG	ATATTGACAT	AAATGCAACA	1080
	GTTCGCTAGG	ACCTTGAAGT	TCCAATGACC	CGAGAAAAC	ATGAAAAGCT	ATGGAAATTT	1140
	ATAGCATACT	ACAGTGAAGT	TCCGTGAAG	CTACACAGAG	AGCTCATGCT	CAGCAAAGAC	1200
25	CCCAGAGTCA	GCTACCAGTA	CAGGCAGGAT	GCTGATGAGG	AAGCTCTTTA	CTACACAGGT	1260
	GTGAGAGCCC	AGATTCTTGC	AGAACCAGAA	TGGGTCTATG	AGCCATCCAT	AGATATCCAG	1320
	CTGAACCCAG	CTCAGAGTAC	GGCCAGAAG	GTGCTACTTT	CCTACTACAC	CCAGTATTCT	1380
	CAAAACAATAT	CCACCAAGAA	TACAAGGCAG	GCTCGGGGCA	GAAGCTGGGT	AATGATTGAG	1440
	CCTAGTGGAG	CTGTGCAAG	AGATCAGACT	GTCCTGGAAG	GGGGTCCATG	CCAGTTGAGC	1500
30	TGCAACGTGA	AAGCTTCTGA	GAGTCCATCT	ATCTTCTGGG	TGCTTCCAGA	TGGCTCCATC	1560
	CTGAAAGCGC	CCATGATGTA	CCAGACAGC	AAGTTCTCCA	TTCTCAGCAG	TGGCTGGCTG	1620
	AGGATCAAGT	CCATGGAGCG	ATCTGACTCA	GGCTTGTACC	AGTGATGCTG	TCAAGTGAGG	1680
	GATGAAATGG	ACCGCATGGT	ATATAGGGTA	CTTGTGCACT	CTCCCTCCAC	TCAGCCAGCC	1740
	GAGAAAGACA	CAGTGACAAT	TGGCAAGAAC	CCAGGGGAGT	CGGTGACATT	GCCTTGCAAT	1800
35	GCTTTAGCAA	TACCCGAAGC	CCACCTTAGC	TGGATTCTTC	CAAAACAGAG	GATAATTAAT	1860
	GATTTGGGCT	ACATCATACA	TGTATACATG	TGCAAAATG	GAATCTTTTC	CATCCCAAAG	1920
	GTCCAACTCA	GATGATGCTG	TTACTACAGA	TGTGTGGCTG	TCAACCAGCA	AGGGGCGAGC	1980
	CATTTTACGG	TGGGAATCAC	AGTGACCAAG	AAAGGGTCTG	GCTTGCCATC	CAAAAGAGGC	2040
	AGACGCGCCG	GTGCAAGGCG	TCTTTCCAGA	GTGAGAGAG	ACATCGTGGA	GGATGAAGGG	2100
40	GGCTCGGGCA	TGGGAGATGA	AGAGAACACT	TCAAGGAGAC	TTCTGCATCC	AAAGGACCAA	2160
	GAGGTGTTCC	TCAAAACAAA	GGATGATGCC	ATCAATGGAG	ACAAGAAAGC	CAAGAAAGGG	2220
	AGAAGAAAGC	TGAACCTCTG	GAAGCATTCT	GAAAAGAAC	CAGAGACCAA	TGTTGCAGAA	2280
	GGTCGAGAG	TGTTTGAATC	TAGACGAAGG	ATAAACATGG	CAAAACAACA	GATTAATCCG	2340
	GAGCGCTGGG	CTGATATTTT	AGCCAAAGTC	CGTGGGAAAA	ATCTCCCTAA	GGGCACAGAA	2400
45	GTACCCCGGT	TGATTAATAC	CACAAGTCTT	CCATCCTTGA	GCCTAGAAGT	CACACCACCT	2460
	TTTCTGCTGT	TTTCTCCCCC	CTCAGCATCT	CCTGTGCAGA	CAGTAACCAAG	TGCTGAAGAA	2520
	TCCTCAGCAG	ATGTACCTCT	ACTTGGTGAA	GAAGAGCAGG	TTTGGGTAC	CATTCTCTCA	2580
	GCCAGCATGG	GGCTAGAACA	CAACCAATAT	GGAGTTATTC	TTGTTGAACC	TGAAGTAACA	2640
	AGCACACCTC	TGGAGGAAGT	TGTTGATGAC	CTTCTGAGA	AGACTGAGGA	GATAACTTCC	2700
50	ACTGAAGGAG	ACCTGAAGGG	GACAGCAGCC	CCTACACTTA	TATCTGAGCC	TTATGAACCA	2760
	TCTCTACTCT	TGCACACATC	AGACACAGTC	TATGAAGAGC	CCACCCATGA	AGAGACGGCA	2820
	ACAGAGGGTT	GGTCTGCAGC	AGATGTTGGA	TCGTCAACAG	AGCCACATCC	CAGTGAGTAT	2880
	GAGCCTCCAT	TGGATGCTGT	CTCCTTGGCT	GAGTCTGAGC	CCATGCAATA	CTTTGACCCA	2940
	GATTGGGAGA	CTAAGTCACA	ACCAGATGAG	GATAAGATGA	AAGAAGACAC	CTTTGCACAC	3000
55	CTTACTCCAA	CCCCCACCAT	CTGGGTTAAT	GACTCCAGTA	CATCAGAGTT	ATTTGAGGAT	3060
	TCTACTATAG	GGGAACCCAG	TGTCCCAGGC	CAATCACATC	TACAAGGACT	GACAGACAAC	3120
	ATCCACCTTG	TGAAGAGTAG	TCTAAGCACT	CAAGACACCT	TACTGATTAA	AAAGGGTATG	3180
	AAAGAGATGT	CTCAGACACT	ACAGGGAGGA	AATATGCTAG	AGGGAGACCC	CACACACTCC	3240
	AGAAAGTTCT	AGAGTGAGGG	CCAAGAGAGC	AAATCCATCA	CTTTGCCCTGA	CTCCACACTG	3300
60	GGTATAATGA	GCATATGTCT	TCCAGTTAAG	AAGCCTCGGG	AAACCAACAGT	TGGTACCCCT	3360
	CTAGACAAAG	ACACCAACAC	AGTAACAACA	ACACCAAGGC	AAAAAGTTGC	TCCGTCACTC	3420
	ACCATGAGCA	CTCACCTTCT	TCGAAGGAGA	CCCAACGGGA	GAAGGAGATT	ACGCCCAAC	3480
	AAATTCCGCC	ACCGGCACAA	GCAAAACCCA	CCCAACACTT	TTGCCCCATC	AGAGACTTTT	3540
	TCTACTCAAC	CAACTCAAGC	ACCTGACATT	AAGATTTCAA	GTCAAGTGGA	GAGTTCTCTG	3600
65	GTTCTTACAG	CTTGGGTGGA	TAAACACAGT	AATACCCCA	AACAGTTGGA	AATGGAGGAG	3660
	AATGCAGAAC	CCACATCCAA	GGGAACACCA	CGGAGAAAAC	ACGGGAAGAG	GCCAAACAAA	3720
	CATCGATATA	CCCCCTTCTAC	AGTGAAGTCA	AGAGCGTCCG	GATCCAAGCC	CAGCCCTTCT	3780
	CCAGAAAATA	AACATAGAAA	CATTGTTACT	CCAGTTTCA	AACTATACT	TTTGCCTAGA	3840
70	ACTGTTTCTC	TGAAAACCTGA	GGGCCCTTAT	GATTCTTTAG	ATTACATGAC	AAACCCAGAG	3900
	AAAATATATT	CATCTTACCC	TAAAGTCCAA	GAGACACTTC	CAGTCACATA	TAAACCCACA	3960
	TCAGATGGAA	AAGAAATTA	GGATGATGTT	GCCACAAATG	TTGACAAACA	TAAAGTGAGC	4020
	ATTTTAGTCA	CTGGTGAATC	AATTACTAAT	GCCATACCAA	CTTCTGCTCT	CTTGGTCTCC	4080
	ACTATGGGAG	AATTTAAGGA	AGAACTCTCT	CCTGTAGGCT	TTCCAGGAAC	TCCAACCTGG	4140
75	AATCCCTCAA	GGAGGGGCCA	GCCTGGGAGG	CTACAGACAG	ACATACCTGT	TACCACTTCT	4200
	GGGGAAAAAT	TTACAGACCC	TCCCCTTCTT	AAAGAGCTTG	AGGATGTGGA	TTTCACTTCC	4260
	GGGTTTTTGT	CCTCTTTGAC	AGTCTCCACA	CCATTTTACC	AGGAAGAAGC	TGGTTCTTCC	4320
	ACAACCTCTC	CAAGCATAAA	AGTGGAGGTG	GCTTCAAGTC	AGGCAGAAAC	CACCACCTTT	4380
	GATCAGAGTA	ATCTTGAAGC	CAGTGTGGCT	ATTCTCTTTT	CTGAAACTAG	ACCACAGAAT	4440
	CACACCCCTA	TGCTGCCCCG	GATGAAGGAG	CCAGCATCCT	OGTCCCCATC	CACAATTCTC	4500
80	ATGTCTTTGG	GACAAACCC	CACCACTAAG	CCAGCACTTC	CCAGTCCAAG	AATATCTCAA	4560
	GCATCTAGAG	ATTCCAAGGA	AAATGTTTTT	TTGAATTATG	TGGGGAATCC	AGAAACAGAA	4620
	GCAACCCCA	ATCAACATGA	AGGAACACAG	CATATGTCAG	GGCCAAATGA	ATTATCAACA	4680
	CCCTCTTCCG	ACCGGGATGC	ATTTAACTTG	TCTACAAAGC	TGGAATTGGA	AAAGCAAGTA	4740
	TTTGGTAGTA	GAGTCTTACC	ACGTGGGCCA	GATAGCCAAC	GCCAGATGGA	AAGAGTTCA	4800
85	GCTTCTCATC	AACTAACACA	AGTCCCTGCC	AAACCCATCC	TACCAACAGC	AACAGTGAGG	4860
	CTACCTGAAA	TGTCCACACA	AAGCGCTTCC	AGATACTTTG	TAACTTCCCA	GTCACTCTGT	4920
	CAGTGGACCA	ACAAACCCGA	AATAACTACA	TATCCTTCTG	GGGCTTTGCC	AGAGAACAAA	4980
	CAGTTTACAA	CTCCAAGATT	ATCAAGTACA	ACAATTCCTC	TCCCATTTGA	CATGTCACAA	5040

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

CCCAGCATTC CTAGTAAATT TACTGACCGA AGAAGTGAAC AATTCATGG TTAATCCAAA 5100
 GTGTTTGGAA ATAAACAACAT CCTGAGGCA AGAAACCCAG TTGGAAAGCC TCCAGTCCA 5160
 AGAATTCCTC ATTATTCCAA TGGAAAGACT CCTTCTTTTA CCAACAAGAC TCTTCTTTT 5220
 CCACAGTGGG GAGTACCCCG GAGACCCAG ATACCCACTT CTCTGCCCC AGTAATGAGA 5280
 GAGAGAAAAG TTATTCAGG TTCTACAAC AGGATACATT CCAATAGCAC CTTCATCTG 5340
 GACTTTGGCC CTCGGGCACC TCCGTTGTTG CACACTCCGC AGACCAAGGG ATCACCCTCA 5400
 ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC CTTTATAACA 5460
 TCTTCTGTCC AGTCCCTCAGG AAGCTTCCAC CAGAGCAGCT CAAAGTTCTT TGCAGGAGGA 5520
 CTCTCTGATC CCAAAATCTG GTCTCTTGGG GAAAAGCCCC AAATCTCTAC CAAAGTCCCA 5580
 CAGACTGTGT CCGTCAACCCG TGAGACAGAC ACTGTGTTC CCTGTGAGGC AACAGGAAAA 5640
 CCAAAGCCTT TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC 5700
 AGGATACAAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA GGTTCAGTA 5760
 CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC ACGGCCTGGA CAGGATGGTG 5820
 GTCTTGCTTT CGGTACCCGT GCAGCAACCT CAAATCCTAG CTTCCCACTA CCAGGACGTC 5880
 ACTGTCTACC TGGGAGACAC CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCC 5940
 CAAATTTCTT GGAATCTTCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC CGTGAGAGC 6000
 CGCATCACCC TGCACGAAAA CCGGACCCCT TCCATCAAGG AGGCGTCTCT CTCAGACAGA 6060
 GCGGTCTATA AGTGGGTGGC CAGCAATGCA GCGGGGGCGG ACAGCCTGGC CATCCGCTG 6120
 CAGTGGCGCG CACTGCCCCG GGTATCTCAC CAGGAGAAGC TGGAGAACAT CTGCTGCC 6180
 CCGGGGCTCA GCATTACAT TCACTGCACT GCCAAGGCTG CGCCCCCTGC CAGCGTGGC 6240
 TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG GAATCTGTT 6300
 GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCG CCAAGGACAG CGGGCGCTAT 6360
 GAGTGGGTGG CCGCCAACTT GGTAGGCTCC GCGCGCAGGA CGGTGCAGCT GAACTGTCAG 6420
 CGTGCAGCAG CCAACGCGCG CATCACGGGC ACCTCCCCCG GAGGACGGA CGTCAGGTAC 6480
 GGAGGAACCC TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCGCG CATCTCTGG 6540
 AGGCTGCGGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG AATCAAGGTG 6600
 TTTGCCAATG GGACCTGTGT GGTGAATCA GTGACGGACA AAGATGCGCG AGATTACCTG 6660
 TGGTAGCTC GAAATAAGGT TGGTAGTAC TACGTGGTGC TCAAGTGGG TGTGGTGATG 6720
 AAACCGGCA AGATTGAACA CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGTGAC 6780
 CTGAAGTGG ACTGTGTGGC CACGGGCTT CCAATCCCG AGATCTCTG GAGCCTCCCA 6840
 GAGCGGAGTC TGGTGAATC CTTCATGAC TCGGATGACA GCGGTGGACG CACCAAGCGC 6900
 TATGTGCTCT TCAACAATG GACACTCTAC TTTAACGAA TGGGGATGAG GAGGAAAGGA 6960
 GACTACACCT GCTTTGCTGA AAATCAGGTC GGGAAAGGAG AGATGAGAGT CAGAGTCAAG 7020
 GTGGTGACAG CGCCCGCCAC CATCCGGAAC AAGACTTACT TGGCGTTCA GGTGCCCTAT 7080
 GGAGAGCTGG CTGACTGAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCCA GGTGACTGG 7140
 TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT ATACCAAGAT 7200
 GGCACCTCTC TTATTCAGAA AGCCAGCGT TCTGACAGCG GCAACTACAC CTGCTGGTC 7260
 AGGAACAGCG CGGAGAGGAG TAGGAAGACG GTGTGGATT CAGTCAAGT CCAGCCACCC 7320
 AAGATCAACG GTAACCCCAA CCCCATCACC ACCGTGGGG AGATAGCAGC CGGGGGCAGT 7380
 CGGAAGCTGA TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGGT ATGGGCTTTT 7440
 CCGAGGGTGG TGGTCTGCCC AGCTCCATAC TATGGAAGCC GGATCACTGT CCATGGCAAC 7500
 GGTTCCTGGG ACATCAGGAG TTTGAGGAGG AGCGACTCG TCCAGTGGT ATGCATGGCA 7560
 CGCAACGAGG GAGGGGAGGC GAGGTTGATC GTGCAGTCA CTGTCTTGA GCCCATGGAG 7620
 AAACCCATCT TCCAGGACCC GATCAGGAG AAGATCAAGC CCATGGCGGG CCACACCATC 7680
 AGCCTCAACT GCTCTGAGG GGGGACCCCG ACACCCAGCG TGGTGTGGGT CCTTCCCAAT 7740
 GGCACCGATG TGACAGGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC TGAACGGCATG 7800
 CTACACATTA GCGGTCTGCT CTCGGTGGAC GCTGGGGCTT ACCGCTGCTT GGCCCGCAAT 7860
 GCGCTGGGCC ACACGAGAGG GCTGGTCTCC CTGAAGGTGG GACTGAAGCC AGAAGCAAAC 7920
 AAGCAGTATC ATAACCTGGT CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC 7980
 CCTCCCGGGG CTGCGGAGCG ACSTTTCTCC TGGAGCGTCC CCAATGGCAT GCATCTGGAG 8040
 GGCCCCCAAA CCCTGGGAGC CGTTTCTCTT CTGACAAATG GCACCCTCAC GGTTCGTGAG 8100
 GCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG AGACGGAGTA CGGCCCTTG 8160
 GTCACACGAG TCCCGTGAT TGTGATCGCC TATCTCTCCC GGATCACCAG CGAGCCACC 8220
 CCGGTCTATC ACACCCGGCC CGGGAACACC GTGAACTGA ACTGCATGGC TATGGGGATT 8280
 CCAAGAGCTG ACATCACTG GAGGTTACCG GATAAGTCCG ATCTGAAGGC AGGGGTTGAG 8340
 GCTGCTGTGT ATGGAACAG ATTCTTTCAC CCCCAGGAT CACTGACCAT CCAGCATGCC 8400
 ACACAGAGAG ATGCGCGCTT CTACAAGTGC ATGCAAAAAC ACATCTCGG CAGTGACTCC 8460
 AAAACACTT ACATCCACGT CTTCTGAAAT GTGGATTCCA GAATGATTGC TTAGGAACTG 8520
 ACAACAAGC GGGGTTTGTG AGGGAAGCCA GGTGGGGGAA TAGGAGCTCT TAAATATGT 8580
 GTACAGTGC ATGGTGGGCT CTGGTGGGTT TCAAGTTGAG GTTGATCTG ATCTACAATT 8640
 GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTGTCT GAGACACTT 8700
 CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAAGGTG TCTGTGCTCT GACTGCAATT 8760
 TTTCTTCTTT TGCAATGCCC ACTCGACTGC CTTTATAAGC GTCCATAGGA TATCTGAGGA 8820
 ACATTATCAT AAAATAAGCC ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA 8880
 TCACCTAGTT AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCAGAT TGACAAGTCA 8940
 TCTTTCAGTT ATTTCTCTG TCACTTCAA ACTCCAGCTT GCCCAATAAG GATTTAGAAC 9000
 CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT TACATACATA CAGCTACCAT 9060
 TTTATATGAA AARAGAAAAA CATTTCTTCC TGGAACTCAC TTTTATATA ATGTTTTATA 9120
 TATATATTTT TTCTTTTCAA ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT 9180
 ATTAAAAATTA ATAAATATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA 9240
 ATATAATTTT AAAAATTTT TCTCCAACCT CCTTCAAATT CAGTCAACAC TGTATATTA 9300
 CCTTCTCCAG GAACCTTCCA GTGGGGAAGG CTGCGATATT AGATTTCTCT GTATGCAAG 9360
 TTTTGTGTA AAGCTGTGCT CAGAGGAGGT GAGAGGAGAG GAAGGAGAAA ACTGCATCAT 9420
 AACTTTACAG AATTAATCT AGAGTCTTCC CCGAAAAGCC CAGAAACTTC TCTGAGTAT 9480
 CTGCTGTGC CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC 9540
 CCAATGAATA TACAGACCT GTTATTCCA TGACTGCTTT ACTGTATTTT TAAGTCAAT 9600
 ATACTGTACA TTTGATAATA AAATAATATT TCCCCAAAA AAAAA

Seq ID NO: 417 Protein sequence
 Protein Accession #: NP_056234.1

1 11 21 31 41 51
 | | | | |
 85 MPKRAHWGAL SVVLLILLWGH PRVALACPHP CACYVPSEVH CTFRSLASVP AGIARRVERI 60
 NLGENSIQAL SETSPAGLTK LELLMIHNE IPSIPDGALR DLSSLVQPKF SYNKLRVITG 120
 QTLQGLSNLM RLHIDHNKIE PIHPQAPNGL TSLRLHLLEG NLLHQLHPST FSTPTFLDYF 180
 RLSTIRHLYL AENMVRTLPA SMLRNMPLLE NLYLQGNFWT CDCEMRWFLE WDAKSRGILK 240

	CKKDRAYEGG	QLCAMCFSPK	KLYKHEIHL	KDMTCLKPSI	ESPLRQNRSR	SIEEEQEQUE	300
	DGGSQILILEK	PQLPQWSISL	NMTDEHGMV	NLVCIDIKPM	DVYKIHNLQ	DPPDIDINAT	360
	VALDFECMT	RENYEKLWKL	IAYSEVPVK	LHRELMLSKD	PRVSYQYRQD	ADDEALYYTG	420
5	VRAQILAEPE	WWMQPSIDIQ	LNRRQSTARK	VLLSYTYQYS	QTISTKDTQ	ARGRSWVMIE	480
	PSGAVQRDQT	VLEGGPCQLS	CNVKASESPS	IFWVLPDGS	LKAPMDPDPS	KFSILSSGWL	540
	RKSMESPSDS	GLYQCIQVVR	DEMDSMVYVR	LVQSPSTQPA	EKDTVTIGKN	PGESVTLPCN	600
	ALAIPEAHL	WILPQNRRIIN	DLANTSHVYM	LPNGTSLIPK	VQVSDSGYYR	CVAVNQOGAD	660
	HFTVGTITVTK	KGSLGPSKRG	RRPGAKALSR	VREDIVEDEG	SGSGMGDEENT	SRLLHPKQDQ	720
	EVLTKTKDDA	INGDKKAKKG	RRKLKLWKHS	EKEPETNVAE	GRRVFESRRR	INMANKQINP	780
10	ERNADILAKV	RGKNLFPKGT	VPPLIKTTSP	PSLSLEVTPP	FPAVSPPSAS	PVQTVTSABE	840
	SSADVPLLGE	EEHVLGTISS	ASMGLEHNNH	GVILVEPEVT	STPLEEVVDD	LSEKTEEITS	900
	TEGDLKGTA	PTLISEPYEP	SPTLHTLDTV	YEKPTHEETA	TEGWSAADVG	SSPEPTSSEY	960
	EPPLDAVSLA	ESEPMQYFDP	DLETKSQDPE	DKMKEDTFAH	LTPPTIWNV	DSSTSOLFED	1020
	STIGEPGVPG	QSHLQGLTDN	IHLVKSLSLST	QDTLLIKGKM	KEMSTQLQGG	NMLEGDPHTS	1080
15	RSSESEGGES	KSITIPDSTL	GIMSSMSPVK	KPAETTVGTL	LDKDTTIVTT	TPRQKVAPSS	1140
	TMSTHPSRRR	PNRRRLRPN	KFRHRKQTP	PTTFAPSETF	STQPTQAPDI	KISSQVSSSL	1200
	VPTAWNDTVS	WTFKQLEMEK	NAEPTSCKGP	RRKHGKRPNK	HRYPSTVSS	RASGSKFSPS	1260
	PENKRNITVT	PSETILLER	TVSLKTEGPF	DSLDMYMTTR	KIYSSYPKVQ	ETLEVITYKPT	1320
	SDGKEIKDDV	ATNVDKHKSD	ILVTGESITN	AIPTSRSLVS	TMGEFKRESS	PVGFPQTPTW	1380
20	NPSRTAQPRR	LQTDIPVTTT	GENLTDPPIL	KELEDVDFTS	EFLSSLTVST	PFHQEEAGSS	1440
	TTLSSIKVEV	ASSQAEITTL	DQDHLLETTA	ILLSETRPQN	HTPTAARMKE	PASSSPSTIL	1500
	MSLGQTTTTK	PLPSPRISQ	ASRDSKENVF	LNYYGNPETE	ATPVNNEGTO	HMSGPNELST	1560
	PSSDRDAFNL	STKLELEKQV	FGSRSLPRGP	DSQRQDGRVH	ASHQLTRVPA	KPILPLATVR	1620
	LPDMSTQAS	RYFVTSQSPR	HWTNKPEITT	YPSGALPENK	QFTTFLRSST	TIPLPLHMSK	1680
25	PSIPSKPTDR	RTDQNGRYSK	VFGNNNIPEA	RNPVGKPPSP	RIPHYSGNRL	PFTNKTLSF	1740
	PQLGVTRRPQ	IPTSPAPVMR	ERKVIPOGYN	RIHSHSTFHL	DFGPPAPPLL	HTPQTGSPS	1800
	TNLQINPMVS	STQSSISFIT	SSVQSSGSFH	QSSSKFFAGG	PPASKFWSLG	EKPQILTQSP	1860
	QTVSVTAETD	TVPPCEATCK	PKPFVTWTKV	STGALMTFNT	RICRFEVLKN	GTLVIRKQVQ	1920
	QDRGQYMCTA	SNLHGLDRMV	VLLSVTVQPP	QILASHYQDV	TVYLGDTIAM	ECLAKGTPAP	1980
30	QISWIFPDR	RTDQNGRYSK	RITLHENRTL	SIKEASPSDR	GVYKCVASNA	AGADSLAIRL	2040
	HVAALPPVH	QEKLENIPL	PGLSIHIHCT	AKAALPSVR	WVLGDGTQIR	PSQFLHGNLF	2100
	VFPNGTLIIR	NLAPKDSGRY	ECVAANLVGS	ARRTVQLNVQ	RAANARITG	TSPRRTDVRY	2160
	GGTLKLDCA	SGDPWPRILW	RLPSKRMIDA	LFSEFDSRIK	PANGTLVVK	VTDKADGXYL	2220
35	CVARNKVGDD	YVVLKVDVVM	KPAKIEHKEE	NDHKVFYGGD	LKVDVCATGL	PNPEISWSLP	2280
	DGSLVNSPMR	SDDSGGRTKR	YVFNNGTLY	FNEVGMREEG	DYTCFAENQV	GKDEMRVRK	2340
	VVTAPATIRN	KTYLAVQVVP	GDVVTVACEA	KGEPMKVTW	LSPTNKVIFT	SSEKQIYQD	2400
	GTLILQKQR	SDSGNYTCLV	RNSAGEDRKT	VMIHVNQVPP	KINGNPNPIT	TVREIAAGGS	2460
	RKLIDCKAEG	IPTPRVLWAF	PEGVVLPAFY	YGNRITVHGN	GSLLDIRSLR	SDSVQLVCM	2520
40	RNEGGEARLI	VQLTVLEPME	KPIFDPISE	KITAMAGHTI	SLNCSAAGTP	TPSLVWVLPN	2580
	GTDLQSGQQL	ORFYHKADGM	LHISGLSSVD	AGAYRCVARN	AAGHTERLVS	LKVLKPEAN	2640
	KQYHNLVSI	NGETLKLPT	PPGAGQGRFS	WTFPNGMHLE	GPQTLGRVSL	LDNGTLTVRE	2700
	ASVFRDGTIV	CRMETEYGPS	VTSIPVIVIA	YPRITSEPT	PVIYTRPGNT	VKLNCMAMGI	2760
	PKADITWELP	DKSHLKAGVQ	ARLYGNRFLH	PQGSLLTIQHA	TQRDAGFYKC	MAKNILGSDS	2820
45	KTYYIHVF						

Seq ID NO: 418 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..5001

50	1	11	21	31	41	51	
	ATGCCAGGCA	AAAACTAAC	CCGAACAGGC	GCCCCAGCAG	ACTACAGAGT	GATATTGAAG	60
	ACCTCTCAAG	AGGACGAATT	GGATGTACCT	GACGACATCA	CGCTCCGGGT	TATGTTCATCT	120
55	CAGTCTGTGC	TGTGTCTCTG	GGTGGATCCT	GTTCTGGAAA	AACAGAAGAA	AGTTGTITGCA	180
	TCAAGACAGT	ACACCGTGGC	CTATCGAGAG	AAGGGGGAAT	TGGCCAGGTG	GGATTATAAG	240
	CAGATGCGTA	ACAGGCGTGT	GCTGATTGAG	AACCTGATTC	CAGACACTGT	GTATGAATTT	300
	GCAGTCCGTA	TTTCACAGGG	TGAAAGAGAT	GCACAAATGGA	GTACGTCAGT	CTTCCAAAGA	360
	ACACCAGAAAT	CTGCCCTTAC	CACAGCTCCT	GAACACTTGA	ACGCTGGGCC	AGTCAATGGC	420
60	AAACCTACAG	TGTGCTGTGC	ATCTTGGGAT	CGCTTACCAG	AGACTGAGGG	GAAAGTGAAA	480
	GTCTGTCTGC	TGGACACAGG	ACTGTTTCA	GTTTCTCTCT	TCCCAACATC	TGCCAATCA	540
	TTTCAGAATA	CATTCTTTCA	TACGCCCGCG	CTCTCAAAAC	ATTGGAGACA	AAGTCCCTCA	600
	CCTATCTCTG	AGACACTACT	TCTGCCCTGG	TGATGCTGCT	GCAGCCTGGG	GAACGCTATC	660
	TTTTCAAAAT	CGGGGCCACA	AACAGGAGAG	GCCTGGGACC	TCAGTCCAAA	GCCTTCATTG	720
65	TGCTATATCC	AACAAGATG	CAGCTGTACC	CAGAAGGATT	TCAGTTGTCT	AGCTTACCTG	780
	ATCGATATCC	AAACCAACA	AGTTAATAAA	GATCCACAAC	TGGAAGGGAG	TGTTTTTGA	840
	CCATGTTTTC	TTTTCTACTT	CCTCACATTT	ATGCTGGATA	TTGGCGGCTT	TTCTTTCATT	900
	ATGTGCTATG	AAGACCCANN	TGTTTCTTCT	TGACAGGCA	ATTCTTTAAA	ATCTGTTGCA	960
	GCCAGTAAGG	CGGATGTTCA	GCAGAACACG	GAGGACAATG	GGAAAACCCGA	AAAACCTGAG	1020
70	CCTTCTCAC	CTTCTCCAG	AGCTCCAGCT	TCCTCCCAAC	ACCCCTCTGT	GCCTGCTTCT	1080
	CCCCAAGGGA	GAAATGCCAA	GGACCTTCTT	CTTGACTTGA	AGAACAAAAT	ATTGGCTAAT	1140
	GGTGGGGGCG	CCGAAAACCC	CCAGCTTCGC	GCCAAGAAGG	CAGAGGAGCT	GGATCTTCAG	1200
	TCGACAGAAA	TACTGGGGA	GGAGGAGCTG	GTTTCCCGGG	AGGACTCGCC	CATGTCACCC	1260
	TCAGACACCC	AAGACCAGAA	ACGGACCCCT	AGGCCGCCAA	GTAGACACGG	CCACTCGGTG	1320
75	GTTGCTCCCG	GACGACTGCG	AGTGAGGGCC	CGGATGCCAG	CGCTGCCCGG	AAGGGAAGGC	1380
	GTAGATAAGC	CTGGCTTTTC	CTTGCCACAG	CAGCCCGGCC	CAGGGGCGCC	CCCCCTGGCT	1440
	TGCGCTCTCT	CTGCCACCA	CGGTCCACCC	CAGGGCACCT	CTCATGCTCC	TTCCCTGCTT	1500
	GCCAGCTTGA	ATGACAACGA	CTTGGTGGAC	TCAGACGAAG	ATGAGCGCGC	TGTGGGCTCC	1560
	CTCACCCCA	AGGCCAGCTT	CGCCAGCCCC	TGTCGCCCGG	CGGCCAGTCC	CGGCCAGTCC	1620
	CGGTCCAGCG	TTCTCCGCGA	CAGAAGCTCT	GTGCACCCCG	CGCCAAAGCC	AGCCTCGCGG	1680
80	GGCGGAGGCA	CCCCCAATTC	AGGGGCCGCA	GAGGAAGATT	CCAGTGCTCT	AGCCCCACCC	1740
	TCAAGACTTT	CTCCACCCCA	TGGGGGATCA	TCTCGGCTGC	TGCCACCCCA	GCCACACCTG	1800
	AGCTCTCCAC	TTTCCAGGGG	CGGGAAGGAT	GGTGAGGAGC	CCCCAGCCAC	CAACTCCCAAT	1860
	GCGCATCATC	GATCCACCAT	GTCTCTCTCC	GTCTCTCTCT	ATCTCTGCTC	CAGGACGCGA	1920
85	GTCTCTGAGG	GAGCGGAGGC	TTCTGATGGT	GAAGGCCACG	GTGACGCGCA	TAGGGAAGAC	1980
	GGCGGAGAGC	AGCGGAGGCG	CACGGCCGAG	ACGCTGCGGG	CCCGGCTGCT	CTCTGGACAC	2040
	TTCCATTGTC	TCAGACACAA	ACCCTTTGCT	GCCAACGGGA	GGTCTCCAAG	CAGGTTTCAGC	2100
	ATTGGGCGGG	GACCTCGGCT	GCAGCCCTCC	AGCTCCCCAC	AGTGGACTGT	GCCCTCCCGA	2160

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

```

GCCACCCCA GGGTCCCTC TCACTCTGAT TCCACCCCTA AGCTTAGCTC AGGTATCCAT 2220
GGAGACGAGG AGGATGAGAA GCCGCTTCCT GCCACCGTTG TCAATGACCA CGTGCCTTCC 2280
TCCTCCAGGC AGCCCATCTC CGGGGGCTGG GAGGACTTAA GGAGAAAGCCC GCAGAGAGGG 2340
GCCAGCCTGC ATCGGAAGGA ACCCATCCCA GAGAACCCCA AATCCACAGG GGCAGATACA 2400
CATCTCTCAG GCAAGTACTC CTCCTGSGCC TCCAAGGCTC AGGATGTTCA ACAGAGACAA 2460
GACGCGGACA CGGAGGGTCA TTCTCCCAA GCACAGCCAG GGTCCACAGA CCGCCACGCG 2520
TCCCTGCTCT GTCTCTCCGC AGCAAGGTCA CAGCAGCATC CCAATGTTCC CAGAAGGATG 2580
ACACCCGCGC GGGCCCCAGA ACAGCAGCCC CCTCCTCCCG TCGCCACGTC CCAGCACACC 2640
CGGGACCCCG AGAGCAGAGA CGCGGGTCGG TCACCTTCCC AGCCACAGGT CTCACTGACC 2700
CAGGCGGCGG GGGCCCGCCC CACGTGCGAG GGCCTGCTCC ACTCCTCTCT GGACCTTAC 2760
ACGGCAGAGT CCAGAGGGAT GCTCCCCACG GCCCTCCAGA ACCAGGACGA GGATGCCCGG 2820
GGCAGCTACG ACCAGACAGC CACAGAAGTC GAGGCCCAGG ATGTGCGGGC CCGCGCGCAC 2880
GCCGCGCGCG CCAGAGGAGG AGCTGCGTCC CTTCCCAAGC ACCAGCAGGT GGAGTCTCCC 2940
ACAGGCGCAG GGCAGGTGGG CGACACAGG TCCAGCGCG GACATGCGGC CTCCCCCGCC 3000
AGGCCACGCC GACCCGCGCG CCCCAGTCC CGCGCCCGGG TCCCAGCAG GGCAGCGCG 3060
GGGAAGTCGG AGCTCCTTC CAAGCGGCCC CTGTCTTCCA AGTCCCAGCA GTCGGTCTCA 3120
GCCGAGGACG AGGAGGAGGA GGAACGCGGG TTTTAAAG GCGGAAAGA AGACCTTCTG 3180
TCTTCTCTG TGCCAAAGTG GCCCTCTTCC TCCACTCCCA GGGGCGGCAA AGACGCGGAT 3240
GGGAGCCTCG CCAAGGAAGA GAGGAGCCTT GCCATCGCGC TTGCCCCCTG CGGAGGGAGC 3300
CTGGCTCTG TGAAGCGACC TCTCCCCCA CCTCCAGGCA GCTCCCCAG GGCCTCCCAC 3360
GTCCCTTCCC GACCGCGGCC TCGCAGCGCT GCCACCGTGA GCCCGCTGCG GGGCACCACC 3420
CCCTGGCCCG GGTACACCA CGCGGCCCVV CCTGGCCACT TCTCCACCAC CCGATGCTG 3480
TCCTTGGCG AGAGGATGAT GCATGCCAGA TTCCGTAAAC CTCTCTCCCG ACAGCTGCC 3540
AGACCTCTT ACAGCAAGG TTATAATGGC AGACCAATG TAGAAGGAA AGTCTTCTCT 3600
GGTAGTAATG GAAAACCGAA TGGACAGAGA ATTATCAATG GCCCTCAAGG AACAAAGTGG 3660
GTTGTGGACC TTGATCGTGG GTTAGTATTG AATGCAAGG GAAGGTACCT CCAAGATTCA 3720
CATGGAATTC CTCTTGGGAT TAACTAGGA GGAGATGGTC GAACCATTTG AGATCTGAA 3780
GGGACCCCGG TGGTGAAGTC TGACGGCTTC CCACTCTTTG GGCAGGGGCG ACATGGCACA 3840
CCTCTGGCCA ATGCCAAGA TAAGCCAATT TTGAGTCTTG GAGGAAAGCC GCTGGTGGGC 3900
TTGGAGGTCA TCAAAAAAAC CACCCATCCC CCTACCACTA CCATGCAGCC CACCACTACT 3960
ACGACGCCCC TGCTTACCAC TACAACCCCG AGGCCACCA CTGCCACCAC CATGCAGCCC 4020
ACCACTACTA CGACGCCCCC GCCTACCACT ACACCGAGGC CCACCACTGC CACCAACCCG 4080
CGCAGGACCA CCGAGCGTCC AACAAACCA GTCCGAACCA CTACGCGGAC AACCAACACC 4140
ACCAACCCCA AACCCACCAC TCCCATCCCC ACCTGTCCCC CTGGGACCTT GGAACGCGAC 4200
GACGATGATG GCAACCTGAT AATGAGCTCC AATGGGATCC CAGAGTGCTA CGCTGAAGAA 4260
GATGAGTTCT CAGGCTTGA GACTGACACT GCAGTACCTA CGGAAGAGGC CTACGTTATA 4320
TATGATGAAG ATTATGAATT TGAGACGTCA AGGCCACCAA CCACCACTGA GCCTTCGAAC 4380
ACTGCTACCA CACGAGGGT GATCCAGAG GAAGCGGCA TCAGTTCTCT CTCTGAAGAA 4440
GAATTTGATC TGGCTGGAAG GAAACGATT GTTGCTCTCT ACCTGACGTA CCTAAATAAA 4500
GACCCATCAG CCGCGGAGC TCTGACTGAT GCCTGGATC ACTTCCAAGT GGCAGCGCTG 4560
GATGAAATCA TCCCAATGA CCTGAAGAAG AGTGATCTGC CTCCCCAGCA TGCTCCCCGC 4620
AACATCAGCG TGGTGGCTGC GGAAGGTTGC CACTCATTG TCATGTGGA TTGGGACAAA 4680
GCCACCCAG GAGATTGGT CACAGTTAT TTGTTTACA GTGCATCTA TGAAGATTTC 4740
ATCAGGAACA AGTTTTCAC TCAAGTTCA TCAGTAACCT ACTTGCCCAT TGAGAACCTA 4800
AAGCCCAACA CGAGGTATTA TTTTAAAGT CAAGCAAAA ATCCTCATGG CTACGGACCT 4860
ATCAGCCCTT CGTCTCATT TGTCACCGAA TCAGATAATC CTCTGCTTGT TGTGAGGCCC 4920
CCAGGCGGTG AGCTTACTGG ATCCATTGG CTTTCAAAAC TGATCCAGC TACACGACT 4980
GCCATGGAGC GCAATATGTC AAGCGCACGT GGTATCGAAA GTTCGTGGGA GTTGTCTCT 5040
GTAATTCAC TGGATATAA ATCTACCTCA GTGACAACT GAAAGATACA TTCTACAGCA 5100
TTGGAGACAG CTGGGGAAG GGTGAAGACC ATTGCCAATT TGTGGATTCA CACCTTGATG 5160
GAAGAACAGG GCCTCAGTCC TATGTAGAAG CCCTCCCTAC TATTCAAGGC TACTATCGCC 5220
AGTATCGTCA GGAGCTGTC AGGTTGGGA ACATCGGCTT CGGAACCCCT TACTACTATG 5280
TGGGCTGGTA CAGATGTGGG GTCTCCATCC CTGGAAGTG GTAATCAG GACGCTCATG 5340
CTGCAAGCTT GCCTGCCCA GCCCCACCAA CTAAGTGGCA CTAGGGGCTG TGAGCAAGA 5400
CAGCCAGACT GCTCAGCCCC GCTGCCAGG GTGCACAGAT GGCACACTGGC 5460
CATTCTGGTC ATCTCAGTCT GGAACCTCAG CCCACTTCTT GGCCTGGACA ATGAACAGGA 5520
TTCACTTCTG CTGTAACTT TGCTTCTCTA CTTTCTTTT TTTGTTTGA ATAGCACATC 5580
CCAGAGACAT CAGAAACAG CAACTGATTC AGTGTGATTT CCCAGACTTT TTAGGCATGA 5640
AATTCCGACA CTTCAGTATT TCCAGGAATA GCATATGCAC GCTGTTCTTG CTTCATGGA 5700
TGCTACATGC TTCTGTTTTT TCTCATTTTG GATTTCTCCA AAACCTAAGT AATTAAAGT 5760
TCAGGTCCTT TTGATGCGAG TAGAAAGGAA TTATTAAGAA CACCACCAA GAAATAAAT 5820
ATATCTACT TGAAATTTAC TCTATGGACT TACCACTGC TAGAATAAAT GTATCAAATC 5880
TTATTGTAA ATTCTCAATT TTGATATATA TATGTATATA TGCATATACA TATCCACT 5940
TGCTGCAAG AATATTGATT AAAATTGCTA AATTGTACT TGTTCACCAA AAAAAAAA 6000
AAAAAA
  
```

Seq ID NO: 419 Protein sequence
Protein Accession #: Eos sequence

70
 75
 80
 85

```

1 11 21 31 41 51
MPGTLKTRTG APADYRVILK TSQDELDVDF DDISVRVMSS QSVLVSWVDP VLEKQKKVVA 60
SRQYTVRYRE KGELARWDYK QIANRRVLIE NLIPDTVYEF AVRISQGERD GKNSTSVFQR 120
TPESAPTTAP ENLNVVPVNG KPTVVAASWD ALPETEGKVK VCLLDLGLFS VSSPQPSAKS 180
FQNTFFHTPR LSNHLEQSPS PILETLLEPW WMVCSLGNAI FSKSGPQTGE AWDLTTPKPSL 240
SLCQBCSCT QKDFSCLAYL IDIQTKQVNE DPQLEGSVFG PCFLFYPLTF MLDIGGFSFI 300
MCYEDPVSSL TGNLSKSVAA SKADVQNTTE DNGKPEKPEP SSPSPRAPAS SQHPSVPASP 360
QGRNAKDLLE DLKNKILANG GAPRKPQLRA KKAELDLQ9 TEITGEEELG SREDSPMSPS 420
DTQDQKRLR PPSRHHGSVV APGRTAVRAR MPALPRREGV DKPGPSLATQ PRPGAPPSAS 480
ASPAHASTQ GTSRHSPLPA SLNDNDLVDS DEDERAVGSL HPKGAPQPR PALSPSRQSP 540
SSVLRRSSV HPQAPASPA RRTPHSGAE EDSSASAPPS RLSPPHGGSS RLLPTQPHLS 600
SPLSKGKGDG EDAPATNSNA PSRSTMS9V SSHLSSRTQV SEGAESDGE SHGDGREDG 660
GRQAEATAQT LRARPASGHF HLLRHKPFPA NGRSPSPPSI GRGPRLQPS SPQSTVPSRA 720
HPRVPSHSDS HPLKSSG1HG DEEDEKPLPA TVVNDHVPSS SRQPI9RGE DLRRSPQRGA 780
SLRKEPIPE NPKESTADTH PQGKSSSLAS KAQDVQOSTD ADTEGHSFKA QPGSTDRHAS 840
PARPPAARSQ QHPSVPRRMT PGRAPEQPP PFVATSQHEP GPQSRDAGRS PQQRLSLTQ 900
AGRPRPTSQG RSHSSSDPYT ASSRGLMFTA LQNDQEDAQG SYDDSDTEVE AQDVRAPAHA 960
  
```

ARAKEAASL PKHQVESPT GAGAGGDHRS QRGHASPAR PSRPGGPQSR ARVPSRAAPG 1020
KSEPPSKRPL SSKSQSVSA EDEEEDAGF FKGGKEDLLS SSVKWPSSS TPRGGKDADG 1080
SLAKEEREPF IALAPRGMSL APVKRPLPPF PGSSPRASHV PSRPPPSAA TVSPVAGTHP 1140
WPRYTTRAPP GHFSTTPMLS LRQRMHARF RNPLSRQPAR PSYRQGYNGR PNVEGKVLPG 1200
5 SNGKPNQORI INGPQGTQWV VDLDRGLVLN AEGRYLDQSH GNPLRIKLG DGRTIVDLEG 1260
TPVVSPPDGL LFQGGHRGTP LANAQDKPL SLGGKPLVLG EVIKKTHHP TTTMQPTTTT 1320
TLPPTTTTPR PTTATTMQPT TTTTLPPTT PRPTTATRR TTTTRPTTV RTTTRTTTT 1380
TPKPTTPTPT CPPGTLEHRD DDGNLIMSSN GIPECYAED EFSGLETDTA VPTEEAIVY 1440
10 DEDYRFPETSR PFTTTGFSPT ATTPRVIPEE GAISFPPEE FDLAGRKRFF APYVTYLNKD 1500
PSAPCSLTD LDHFPQVDSL EIIPNDLKKS DLPPQHAPRN ITVVAVEGCH SPVIVDNDKA 1560
TPGDLVTGYL VYSASYEDFI RNKFSTQASS VTHLPITENLK PNTRYFKVQ AQNPBGYGP 1620
SPSVSFVTES DNPLLVVRFP GGELSGSHSL SNMIPATRTA MDGNM

Seq ID NO: 420 DNA sequence
Nucleic Acid Accession #: NM_022743
Coding sequence: 128..1237

1 11 21 31 41 51
20 GTGGATTTTA GAGATACCTC CCCTCCTTCT GCTCAGCTGC CTTGCAGTAA TTAACCTCTT 60
TCTCTGCTGC AACACCCCTA CTGTTCTCCG TGTATTGGCT TTTCTGGGCA GCAGGAAGGA 120
AAAGCTGATG CGATGCTCTC AGTGCOCGT CGCCAAATAC TGTAGTGCTA AGTGTACAGAA 180
AAAAGCTTGG CCAGACCACA AGCGGGAATG CAAATGCCTT AAAAGCTGCA AACCCAGATA 240
25 TCCTCCAGAC TCCGTTCCAG TTCTTGGCAG AGTTGTCTTC AAACCTATGG ATGAGCAGCC 300
TTCAGAATCA GAGAAGCTTT ACTCATTTTA TGATCTGGAG TCAAAATATTA ACAAACTGAC 360
TGAAGATAAG AAGAGGGGCC TCAGGCAACT CGTAATGACA TTTCAACATT TCATGAGAGA 420
AGAAATACAG GATGCTCTC AGCTGCCACC TGCCCTTGAC CTTTGTGAAG CCTTGTGAAA 480
AGTGATCTGC AACTCTTCA CCATCTGTAA TGGGAGATG CAGGAAGTTG GTGTTGGCCT 540
ATATCCCACT ATCTCTTTCG TCAATCACAG CTGTGACCCC AACTGTTCGA TTGTGTTCAA 600
30 TGGGCCCCAC CTCTTACTGC GAGCAGTCCG AGACATCGAG GTGGGAGAGG AGCTCACCAT 660
CTGCTACCTG GATATGCTGA TGACCACTGA GGAGCGCCGG AAGCAGCTGA GGGACCACTA 720
CTGCTTTGAA TGTGACTGTT TCCGTTGCCA AACCCAGGAC AAGGATGCTG ATATGCTAAC 780
TGGTGATGAG CAAGTATGGA AGGAAGTTCA AGAATCCCTG AAAAAAATTG AAGAACTGAA 840
GGCACACTGG AAGTGGGAGC AGGTTCTGGC CATGTGCCAG GCGATCATAA GCAGCAATTG 900
35 TGAACGGCTT CCGGATATCA ACATCTACCA GCTGAAGGTG CTCGACTGCG CCATGGATGC 960
CTGCATCAAC CTGCGCCTGT TGGAGGAAGC CTTGTTCTAT GGTACTCGGA CCATGGAGCC 1020
ATACAGGATT TTTTTCCTCG GAAGCCATCC GGTGAGAGG GTTCAAGTGA TGAAGTTGG 1080
CAAACTGCAG CTACATCAAG GCATGTTTCC CCAAGCAATG AAGAACTGGA GACTCGCTTT 1140
TGATATTATG AGAGTGACAC ATGGCAGAGA ACACAGCCTG ATTGAAGATT TGATTCTACT 1200
40 TTTAGAAGAA TGGCAGGCCA ACATCAGAGC ATCCTAAGGG AACGCAGTCA GAGGGAATA 1260
CGGCGTGTGT CTTTGTGTA TGCCTTATTG AGGTACACACA CTCTATGCTT TGTATGCTGT 1320
GTGAACCTCT CTTATTGGA ATTCTGTTC GTGTTTGTGT AGGTAAATAA AGGCAGACAT 1380
GGTTTGCAA CCACAAGAA CATTAGTTGT AGAGAAGCAC GATTATAATA AATTCAAAAC 1440
45 ATTTGGTTGA GGATGCCAAA AAAAAAATA AAAAAA

Seq ID NO: 421 Protein sequence
Protein Accession #: NP_073580

1 11 21 31 41 51
50 MRCSQCRVAK YCSAKQKKA WFDHKRECKC LKSCIKPRYP DSVRLGRVV FKLMGAPSE 60
SEKLYSPYDL ESNINKLTED KKEGLRQLVM TPQHFMRBEE QDASQLPPAF DLFEAFKVI 120
CNSPTICNAE MQEVGVGLYP SISLLNHSCD PNCSTIVFNGP HLLLRVARDI EVGEELTICY 180
LDMLMTSEER RKQIRDQYCF ECDPCFRQQT DMDADMLTGD EQVWKEVQES LKKIEELKAH 240
55 WKWEQLVANC QALISSNSER LFDINITYQLK VLDCAMDACI NLGLLEALF YGTRTMEFYR 300
IFPPGSHFVR GVQVMKVGLK QLEQGMFPQA MNLRALAFDI MRVTGREHS LIEDLILILE 360
ECDANIRAS

Seq ID NO: 422 DNA sequence
Nucleic Acid Accession #: NM_003014.2
Coding sequence: 238..648

1 11 21 31 41 51
65 GGGGGGTTGG CGCCCCGAAG GCTGAGAGCT GGGCTGCTC GTGCCCTGTG TGCCAGACGG 60
CGGAGCTCCG CGGCGGAGCC CGGCGGCCCC GCTTGTCTGC CGACTGGAGT TTGGGGGAAG 120
AAACTCTCCT CGGCCCCAGA AGATTTCTTC CTCGGCGAAG GGACAGCGAA AGATGAGGGT 180
GGCAGGAAGA GAAGGCGCTT TCTGTCTGCC GGGGTGCGAG CGCGAGAGGG CAGTGCCATG 240
70 TTCTCTCCA TCCTAGTGGC GCTGTGCTGC TGGCTGCACC TGGCGCTGGG CGTGCGCGCC 300
GCGCCCTGCG AGGCGGTGCG CATCCCTATG TGCCGCGACA TGCCCTGGAA CATCACGCGG 360
ATGCCCAACC ACCTGCACCA CAGCAGCAG GAGAACGCCA TCCTGGCCAT CGAGCAGTAC 420
GAGGAGCTGG TGGACGTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTCTG TGCCATGTAC 480
GCGCCCATTT GCACCTTGA GTTCTCTGAC GACCCTATCA AGCCGTGCAA GTCCGTGTGC 540
75 CAACGCGCGC GCGACGACTG CGAGCCCTTC ATGAAGATGT ACAACACAG CTGGCCCGAA 600
AGCTCGGCTT GCGACGAGCT GCTGTCTAT GACCGTGGCG TGTGCAITTC GCCTGAAGCC 660
ATCGTCAAGG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACCAGA CATGATGGTA 720
CAGGAAGGC CTCTTGATGT TGACTGTAAA CGCCTAAGCC CGGATCGGTG CAAAGTAAA 780
AAGGTGAAGC CAACCTTGGC AACGTATCTC AGCAAAACT ACAGCTATGT TATTATGCC 840
80 AAAATAAAG CTGTGAGAGG GAGTGGCTGC AATGAGGTCA CAACGCTGGT GGATGTAAA 900
GAGATCTTCA AGTCTCATC ACCCATCCCT CGAACTCAAG TCCGCTCAT TACAAATCT 960
TCTTGCCAGT GTCCACACAT CCTGCCCCAT CAAGATGTTT TCATCATGTG TTAGAGTGG 1020
CGTTCAAGGA TGATGCTTCT TGAATAATGC TTAGTTGAAA AATGAGAGA TCAGCTTAGT 1080
AAAAGATCCA TACAGTGGG AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAG 1140
AAGAAACAG CCGGCGCAC CAGTCTAGT AATCCCCCA AACCAAGGG AAAGCTCCT 1200
85 GCTCCCAAC CAGCGAGTCC CAAGAAGAAC ATTAATACTA GGAGTGCCCA GAAGAGACA 1260
AACCCGAAA GAGTGTGAGC TAACTAGTTT CCAAGCGGA GACTTCGAC TTCCTACAG 1320
GATGAGGCTG GGCATTGCCT GGGACAGCCT ATGTAAGGCC ATGTGCCCT TGCCCTAACA 1380

ACTCACTGCA GTGCTCTTCA TAGACACATC TTGCAGCATT TTTCTTAAGG CTATGCTTCA 1440
 GTTTTCTTTT GTAAGCCATC ACAAGCCATA GTGGTAGGTT TGCCCTTTGG TACAGAAGGT 1500
 GAGTTAAAGC TGGTGGAAAA GGCTTATTGC ATTGCATTCA GAGTAACCTG TGTGCATACT 1560
 CTAGAAGAGT AGGGAAAAATA ATGCTTGTTA CAATTGACCC TAATATGTGC ATTGTAAAAAT 1620
 AAATGCCATA TTTCAACAA AACACGTAAT TTTTTCACAG TATGTTTTAT TACCTTTTGA 1680
 TATCTGTGTG TGCATGTGTA GTGATGTTTT AAAATGTGAT GAAAATATAA TGTTTTTAAG 1740
 AAGGAACAGT AGTGGAAATGA ATGTTAAAAA ATCTTTATGT GTTTATGGTC TGCAGAAAGG 1800
 TTTTGTGATG GAAAGGGGAT TTTTGAAGAA ATTAGAGAAG TAGCATATGG AAAATATATA 1860
 TGTGTTTTTT TACCAATGAC TTCAGTTTCT GTTTTATAGT AGAAACTTAA AAACAAAAAT 1920
 AATAATAAAG AAAATAAAT AAAAAGGAGA GGCAGACAAT GTCTGCATTC CTGTTTTTTG 1980
 GTTACCTGAT TTCCATGATC ATGATGCTTC TTGTCAACAC CCTCTTAAGC AGCACCAGAA 2040
 ACAGTGAGTT TGTCTGTACC ATTAGGAGTT AGGTACTAAT TAGTTGGCTA ATGCTCAAGT 2100
 ATTTTATACC CACAAGAGAG GTATGTCACT CATCTTACTT CCCAGGACAT CCACCCCTGAG 2160
 AATAATTGA CAAGCTTAAA AATGGCCTTC ATGTGAGTGC CAAATTTTGT TTTTCTTCAT 2220
 TTAATATATT TCTTTCCTTA AATACATGTG AGAGGAGTTA AATATAAATG TACAGAGAGG 2280
 AAAGTTAGT TCCACCTCTG AAATGAGAAT TACTTGACAG TTGGGATACT TTAATCAGAA 2340
 AAAAAGAAAT TATTTGACAG ATTTTATCAA CAAATTTTCA AATTGTGAC AATTGGAGGC 2400
 ATTTATTTTA AAAACAATT TTATTGGCCT TTTGCTAACA CAGTAAGCAT GTATTTTATA 2460
 AGGCATTCAA TAAATGCACA ACGCCCAAAG GAAATAAAT CCTATCTAAT CCTACTCTCC 2520
 ACTACACAGA GGTATCTACT ATTAGTATTT TGGCATATTA TTCTCCAGGT GTTGTCTTAT 2580
 GCACTTATAA AATGATTGGA ACAAAATAAA CTAGGAACCT GTATACATGT GTTTCATAAC 2640
 TCGCTCTGTT TGTCTGGCCC TTTATTGAGA TAAGTTTTCG TGTCAAGAAA GCAGAAACCA 2700
 TCTATTCTCT AACAGCTGTG TTATATTCCA TAGTATGCAT TACTCAACAA ACTGTTGTGC 2760
 TATTGGATAC TTAGTGGGTT TCTTCACTGA CAATACTGAA TAAACATCTC ACCGGAATTC

Seq ID NO: 423 Protein sequence

Protein Accession #: NP_003005.1.

1 11 21 31 41 51
 MFLSILVALC LMLHLALGVR GAPCEAVRIP MCRHMPWNIT RMPNHLHST QENAILAIEQ 60
 YEELVDVNCV AVLRFFPCAM YAPICTLEFL HDPIKPKSV QQRARDDECP LMKMYNHSWP 120
 ESLACDELPHV YDRGVCSISPE AIVTDLPELV KWIDITPDMM VQERPLDVDC KRLSPDRCKC 180
 KKVKPLLATY LSKNYSYVIH AKIKAVQRSG CNEVTVVDV KEIFKSSSPI PRQVPLITN 240
 SSCQCPHILP HQDVLIMCYE WSRMMLLEN CLVEKWRDL SKRSIQWEER LQEQRRTVQD 300
 KKKTAGRTSR SNPPKFKGKP PPKPASPKK NIKTRSAQRK TNPKRV

Seq ID NO: 424 DNA sequence

Nucleic Acid Accession #: BC010423

Coding sequence: 248..1780

1 11 21 31 41 51
 CACAGCGTGG GAAGCAGCTC TGGGGGAGCT CGGAGCTCCC GATCACGGCT TCTTGGGGGT 60
 AGCTACGCGT GGGTGTGTAG AACGGGGCCG GGGCTGGGGC TGGGTCCCTT AGTGGAGACC 120
 CAACTGCGAG AGGCAGAAC TCTGCAGCTT CCTGCCTTCT GGGTCAGTTC CTTATTCAAG 180
 TCTGCAGCCG GCTCCCGAGG AGATCTCGST GGAACCTCAG AAACGCTGGG CAGTCTGCCT 240
 TTCAACCATG CCCCTGTCCC TGGGAGCCGA GATGTGGGGG CCTGAGGCCT GGCCTGCTGT 300
 GCTGCTACTG TCGGCATCAT TTACAGGCCG GTGCCCCCGG GGTGAGCTGG AGACCTCAGA 360
 CGTGGTAACT GTGGTCTGGG GCCAGGACGC AAAACTGCCC TGCTTCTACC GAGGGGACTC 420
 CGGCGAGCAA GTGGGGCAAT TGGCATGGGC TCGGGTGGAC GCGGGCGAAG GCGCCACAGG 480
 ACTAGCCTGA CTGCATCCA AATACGGGCT TCATGTGAGC CGGCTTACG AGGGCCGCGT 540
 GGAGCAGCCG CCGCCCCCAC GCACCCCTCT GGACGGCTCA GTGCTCTCTG CCAACGCACT 600
 GCAGGCGGAT GAGGCGGAGT ACGAGTGGCG GGTGAGCACC TTCCCGCGCG GCAGCTTCCA 660
 GGCGCGGCTG GCGCTCCGAG TGCTGGTGCC TCCCTGCCCC TCACTGAATC CTGGTCCAGC 720
 ACTAGAAGAG GGGCAGGGCC TGACCTCTGG AGCCTCTCTG ACAGCTGAGG GCAGCCACAG 780
 CCCCAGCGTG ACCTGGGACA CGGAGGTCAA AGGCACAAAG TCCAGCCGTT CCTTCAAGCA 840
 CTCCCGCTCT GCTGCGGTCA CCTCAGAGTT CCACTTGGTG CCTAGCCGCA GCATGAATGG 900
 GCAGCCACTG ACTTGTGTGG TGTCCCATCC TGGCCTGCTC CAGGACCAAA GGTATCAACC 960
 CATCTTCCAC TTGCTCTTCC TTGCTGAGGC CTCTGTGAGG GGCCTTGAAG ACCAAATCT 1020
 GTGGCATTAT GGCAGAGAAG GAGCTATGCT CAAGTGCTGT AGTGAAGGGC AGCCCCCTCC 1080
 CTCATACAAC TGGACACGCG TGGATGGGCC TCTGCCAGT GGGGTACGAG TGGATGGGGA 1140
 CACTTTGGGC TTTCCCCCAC TGACCACTGA GCACAGGGCG ATCTACGTCT GCCATGTGAG 1200
 CAATGAGTTC TCCTCAAGGG ATTCTCAGGT CACTGTGGAT GTTCTTGACC CCCAGGAAGA 1260
 CTCGCGGAAG CAGGTGGACC TAGTGTGAGC CTCGGTGGTG GTGGTGGGTG TGATCGCCGC 1320
 ACTCTGTGTC TGCTCTCTGG TGGTGGTGGT GGTGCTCATG TCCCGATACC ATCGCGCAAA 1380
 GGCCAGCAGG ATGACCCAGA AATATGAGGA GGAGCTGACC CTGACCAAGG AGAACTCCAT 1440
 CCGGAGGCTG CATTCCTATC ACACGGACCC CAGGAGCCAG CCGGAGGAGA GTGTAGGGCT 1500
 GAGAGCCGAG GGGCACCCCT ATAGTCTCAA GGACAACAGT AGTGTCTCTG TGATGAGTGA 1560
 AGAGCCCGAG GGCCCGAGTT ACTCCACGCT GACCAACGTT AGGGAGATAG AAACACAGAC 1620
 TGAATCTGCT TCTCCAGGCT CTGGGCGGGC CGAGGAGGAG GAAGATCAGG ATGAAGGCAT 1680
 CAAACAGGCC ATGAACCAAT TTGTTTCAAG GAATGGGACC CTACCGGCCA AGCCCAAGGG 1740
 CAATGGCATC TACATCAATG GGCGGGGACA CCTGGTCTGA CCCAGGCTCG CCTCCCTTCC 1800
 CTAGGCGCTG TCTCTTCTGT TGACATGGGA GATTTTAGCT CATCTTGGGG GCCTCCTTAA 1860
 ACACCCCAT TCTTGGCGGA AGATGCTCCC CATCCCACTG ACTGCTTGAC CTTTACCTCC 1920
 AACCTTCTG TTATCGGGA GGGCTCCACC AATTGAGTCT CTCCACCAT GCATGACAGT 1980
 CACTGTGTGT GTGATGTGT GCTGTGTGTA GTGTTGACTG ACTGTGTGTG TGTGGAGGGG 2040
 TGAATGTCCG TGGAGGGGGT ACTGTGTCCG TGGTGTGTAT TATGCTGTCA TATCAGAGTC 2100
 AAGTGAAGCT TGTGTATGT GCCACGGGAT TTGAGTGGTT GCGTGGGCAA CACTGTGAGG 2160
 GTTTGGCGTG TGTGTATGT GGCTGTGTGT GACCTCTGCC TGAAGAAAGCA GGTATTTTCT 2220
 CAGACCCAG AGCAGTATTA ATGATGCAGA GGTGAGGGA GAGAGGTGGA GACTGTGGCT 2280
 CAGACCCAGG TGTGCGGGCA TAGCTGGAGC TGGAACTGCG CTCCGCTGTG AGGGAACCTG 2340
 TCTCCTACCA CTTCGGAGCC ATGGGGGCAA GTGTGAAGCA GCCAGTCCCT GGGTCAGCCA 2400
 GAGGCTTGA CTGTTACAGA AGCCCTCTGC CCTCTGGTGG CCTCTGGGCC TGCTGCATGT 2460
 ACATATTTTC TGTAAATATA CATGCGCCGG GAGCTTCTTG CAGGAATACT GCTCGAATC 2520
 ACTTTTAATT TTTTCTTTT TTTTCTTCTG CCCTTTCCAT TAGTTGTATT TTTTATTTAT 2580
 TTTTATTTT ATTTTCTTTT AGAGTTTGAG TCCAGCCTGG ACGATATAGC CAGACCCCTG 2640

CTGTAAAAAA ACCAAACCC AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 425 Protein sequence
Protein Accession #: AAH10423

5

10

15

1	11	21	31	41	51	
MPLSLGAEMW	GPEANWLLLL	LLASFTGRCP	AGELETSDDV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAMARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRVS	TFPAGSFQAR	LRRLVLPPL	PSLNPGLALE	BQQLTLAAS	CTABGSPAPS	180
VTWDTVEVKT	TSSRSFKHSR	SAAVTSEFHL	VPSRSMNGQP	LTCVVSHPLG	LQDQRIITHL	240
HVSFLAEASV	RGLEDQNLWH	IGREGAMLC	LSEGQPPPSY	NWTRLDGPLE	SGVRVDGDTL	300
GFPPPLTTEHS	GIYVCHVSNE	FSSRDSQVTV	DVLDPQEDSG	KQVDLVASV	VVVGVIAALL	360
FCLLVVVVVL	MSRYHRRKQA	QMTQKYEEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
BGHPDSLKDN	SSCSVMSEEP	EGRSYSTLT	VREIETQTEL	LSPGSGRAEE	EEDQDEGIQ	480
AMNHPVQENG	TLRAKPTGNG	IYINGRHLV				

Seq ID NO: 426 DNA sequence
Nucleic Acid Accession #: NM_003474.2
Coding sequence: 37..3036

20

25

30

35

40

45

50

55

60

65

70

75

80

85

1	11	21	31	41	51	
CACTAACGCT	CTTCTGATGC	CCCGGGCCAA	CTCGGACAGT	TTGCTCATTT	ATTGCAACGG	60
TCAAGGCTGG	CTTGTGCCAG	AACGGCGCGC	GCGCGACGCA	CGCACACACA	CGGGGGGAAA	120
CTTTTTTAAA	AATGAAAGGC	TAGAAGAGCT	CAGCGGCGCG	GCGGGCCGTG	CGCGAGGGCT	180
CCGGAGCTGA	CTCGCGAGG	CAGGAAATCC	CTCGGTGCG	GACGCCCGCG	CCCGCTCGGC	240
GCCCGGCTGG	GATGGTGCAG	CGCTGCGCGC	CGGGCCCGAG	AGCTGTGCA	CTGAAGGCCG	300
GCGACGATGG	CAGCGCGCCC	GCTGCCCGTG	TCCCCCGCCC	GCGCCCTCCT	GCTCGCCCTG	360
GCCGGTGTCT	TGCTCGCGCC	CTGCGAGGCC	CGAGGGGTGA	GCTTATGGAA	CGAAGGAAGA	420
GCTGATGAAG	TTGTCACTGC	CTCTGTTCGG	AGTGGGGACC	TCTGGATCCC	AGTGAAGAGC	480
TTGACTTCCA	AGAAATCATCC	AGAAAGTCTG	AATATTCGAC	TACAAACGGG	AAGCAAAGAA	540
CTGATCATAA	ATCTGAAAG	AAATGAAGGT	CTCATTGCCA	GCAGTTTCAC	GGAAACCCAC	600
TATCTGCAAG	ACGGTACTGA	TGCTCCCTTC	GCTCGAAAT	ACACGTAAT	TCTGGTTCAC	660
TGTTACTACC	ATGGACATGT	ACGGGGATAT	TCTGATTGAG	CAGTCAGTCT	CAGCACGTGT	720
TCTGGTCTCA	GGGGACTTAT	TGTGTTTGAA	AATGAAAGCT	ATGCTTAGA	ACCAATGAAA	780
AGTCAACCA	ACAGATACAA	ACTCTTCCCA	GCGAAGAAGC	TGAAAAGCGT	CCGGGGATCA	840
TGTGATCAC	ATCAACAAC	ACCAAACTTC	GCTGCAAGA	ATGTTTTC	ACCACTCTCT	900
CAGACATGGG	CAAGAAGGCA	TAAAAGAGAG	ACCTCAAGG	CAACTAAGTA	TGTGAGCTG	960
GTGATGCTGG	CAGACAACCG	AGAGTTTCAG	AGGCAAGGAA	AAGATCTGGA	AAAAATTAA	1020
CAGGATTA	TAGAGATTGC	TAATCAAGTT	GACAAGTTT	ACAGACCACT	GAACATTCGG	1080
ATGCTGTTGG	TAGGCTGGA	AGTGTGGAAT	GACATGGACA	AATGCTCTGT	AAGTCAGGAC	1140
CCATTCACCA	GCCTCCATGA	ATTCTGGAC	TGGAGGAAGA	TGAAGCTTCT	ACCTGCGAAA	1200
TCCCATGACA	ATGCGCAGCT	TGTCAGTGGG	GTTTATTTCC	AAGGGACCAC	CATCGGCATG	1260
GCCCCAATCC	TGAGCATGTG	CACGGCAGAC	CAGTCTGGGG	GAATTGTGAT	GGACCATTTA	1320
GACAATCCCC	TTGGTGACGC	CGTGACCTTG	GCACATGAGC	TGGGCCACAA	TTTCGGGATG	1380
AATCATGGA	CATGTGCGG	GGGCTGTAGC	TGTCAAATGG	CGGTTGAGAA	AGGAGGCTGC	1440
ATCATGAACG	CTTCCACCGG	GTACCCATTT	CCCATGTTGT	TCAGCAGTTG	CAGCAGGAAG	1500
GACTGTGGAA	CCAGGCTGGA	GAAAGGAATG	GGGGTGTGCC	TGTTTAACTT	GCGGGAAGTC	1560
AGGGAGTCTT	TGCGGGGCCA	GAAGTGTGGG	AACAGATTG	TGGAAGAAGG	AGAGGAGTGT	1620
GACTGTGGGG	AGGCAAGAGA	ATGTATGAAT	CGCTGCTGCA	ATGCCACCAC	CTGTACCCCTG	1680
AAGCGGAGCG	GTGTGTGCGC	ACATGGGCTG	TGCTGTGAAG	ACTGCCAGCT	GAAGCCTGCA	1740
GGAAACAGCT	GCAGGGACTC	CAGCAACTCC	TGTGACCTCC	CAGAGTTCTG	CACAGGGGCC	1800
AGCCCTCACT	GCACCACCAA	CGTGATCTTG	CACGATGGGC	ACTCATGTCA	GGATGTGGAC	1860
GGCTACTGCT	ACAAATGGCT	CTGCCAGACT	CACGAGCAGC	AGTGTGTGAC	ACTCTGGGGA	1920
CCAGGTGCTA	AACTGCCCCC	TGGGATCTGC	TTTGAGAGAG	TCAATTCTGC	AGGTGATCCT	1980
TATGGCAACT	GTGGCAAGT	CTCGAAGAGT	TCCTTTGCCA	AATGCGAGAT	GAGAGATGCT	2040
AAATGTGGAA	AAATCCAGTG	TCAAGGAGGT	GCCAGCGCGC	CAGTCATTGG	TACCAATGCC	2100
GTTTCCATAG	AAACAACAT	CCCCCTGAG	CAAGGAGGCC	GGATTCTGTG	CCGGGGGACC	2160
CACGTGACT	TGGCGGATGA	CATGCCGAC	CCAGGGCTTG	TGCTTGCAAG	CACAAAGTGT	2220
GCAGATGGA	AAATCTGCCT	GAATCGTCAA	TGTCAAAATA	TTAGTGTCTT	TGGGGTTTAC	2280
GAGTGTGCAA	TGCAGTGCCA	CGGAGAGGG	GTGTGCAACA	ACAGGAAGAA	CTGCCACTGC	2340
GAGGCCCACT	GGGCACCTCC	CTTCTGTGAC	AAGTTTGGCT	TTGGAGGAAG	CACAGACAGC	2400
GGCCCCATCC	GGCAAGCAGA	TAAACCAAGT	TTAACCATAG	GAATTCTGGT	GACCATCTGT	2460
TGCTTCTTGT	CTGCCGGAAT	TGTGGTTTAT	CTCAAAAGGA	AGACCTTGAT	ACGACTGCTG	2520
TTTCAAAATA	AGAGACCAC	CATTGAAAAA	CTAAGGTGTG	TGCGCCCTTC	CCGGCCACCC	2580
CGTGGCTTCC	AACCTGTGCA	GGCTCACCTC	GGCCACCTTG	GAAGAGGCTT	GATGAGGAAG	2640
CGGCCAGATT	CCTACCCACC	GAAGGACAAT	CCCAGGAGAT	TGCTGAGTGC	TCAGAAATGT	2700
GACATCAGCA	GACCCCTCAA	CGGCCTGAAT	GTCCCTCAGC	CCCAGTCAAC	TCAGCGAGTG	2760
CTTCTCTCCC	TCCACCGGGC	CCCACTGTCA	CCTAGCGTCC	CTGCCAGACC	CCTGCCAGCC	2820
AAGCCTGCAC	TTAGGCGAGC	CCAGGGGACC	TGTAAGCCAA	ACCCCTCTCA	GAAGCCTCTG	2880
CCTGCAGATC	CTCTGGCCAG	AACAACCTGG	CTCACTCATG	CCTTGGCCAG	GACCCAGGGA	2940
CAATGGGAGA	CTGGGCTCCG	CCTGGCACCC	CTCAGACCTG	CTCCACAATA	TCCACACCAA	3000
GTGCCAGAT	CCACCCACAC	CGCCTATATT	AAGTGAGAAG	CGACACCTT	TTTTCAACAG	3060
TGAAGACAGA	AGTTTGCACT	ATCTTTCAGC	TCCAGTTGGA	GTTTTTTGTA	CCAACTTTTA	3120
GGATTTTTTT	TAATGTTTAA	AACATCATT	CTATAAGAAC	TTTGAGCTAC	TGCCGTGAGT	3180
GCTGTGCTGT	GCTATGGTGC	TCTGTCTACT	TGCACAGGTA	CTTGTAATTT	ATTAATTTAT	3240
GCAGAAATGT	GATTCAGTGT	CAGTGGCTGT	TAGTAGGCAT	TTTTACCATC	ACTGAGTTTT	3300
CCATGGCAGG	AAGGCTTGT	GTGCTTTTAG	TATTTTAGTG	AACCTGAAAT	ATCTGCTGTG	3360
ATGGGATTC	GGACAGTCT	TGTTTGCTTT	CTGATCAAGG	CCTTATTGGA	AAGCAGTCCC	3420
CCAATACCC	CCAGCTGTGC	TTATGGTACC	AGATGCGAGT	CAAGAGATCC	CAAGTAGAAT	3480
CTCAGTTGAT	TTTCTGGATT	CCCCATCTCA	GGCCAGAGCC	AAGGGGCTTC	AGGTCCAGGC	3540
TGTGTTTGGC	TTTCAAGGAG	GCCCTGTGCC	CCTTGACAAC	TGGCAGGACG	GCTCCCAGGG	3600
ACACCTGGGA	GAAATCTGGC	TTCTGGCCAG	GAAGCTTTGG	TGAGAACCTG	GGTTGCAGAC	3660
AGGAATCTTA	AGGTGTAGCC	ACACAGGAT	AGAGACTGGA	ACACTAGACA	AGCCAGAACT	3720
TGACCTGTAG	CTGACCAAGC	GTGAGCATGT	TTGGAAGGGG	TCTGTAGTGT	CACTCAAGGC	3780
GGTGCTTGAT	AGAAATGCCA	AGCACTCTCT	TTTCTGCTGT	TCCTTCTAG	AGCACTGCCA	3840

5 CCAAGTAGGTT ATTTAGCTTG GGAAAGGTGG TGTTCCTGTA AGAAACCTAC TGCCAGGCA 3900
CTGCAAAACCG CCACCTCCCT ATACTGCTTG GAGCTGAGCA AATCACCACA AACTGTATA 3960
CAATGATCCT GTATTAGAGC AGATGAGGAC TTTCATGGG ACCACAACCTA TTTTCAGATG 4020
TGAACCATTA ACCAGATCTA GTCATCAAG TCTGTTTACT GCAAGGTTCA ACTTATTAAC 4080
AATTAGGCAG ACTCTTTATG CTTGCAAAAA CTACAACCAA TGGAAATGTA TGTTCATGGG 4140
TATAGTTTAT GTCTGCTATC ATTATTGTA GATATTGGAC AAAGAACCTT CTCTATGGGG 4200
CATCCTCTTT TTCCAACCTG GCTGCAGGAA TCTTTAAAG ATGCTTTTAA CAGAGTCTGA 4260
ACCTATTTCT TAAACACTTG CAACCTACCT GTTGAGCATC ACAGAATGTG ATAAGGAAAT 4320
CAACTTGCTT ATCAACTTCC TAAATATTAT GAGATGTGGC TTGGCAGCA TCCCTTGAA 4380
CTCTTCACTC TTCAATGCCC TGACTAGGGA GCCATGTTTC ACAAGGCTTT TAAAGTGACT 4440
AATGGCATGA GAAATACAAA AATACTCAGA TAAGSTAAAA TGCCATGATG CCTCTGCTTT 4500
CTGGACTGGT TTTACATTA GAAGACAATT GACAACAGTT ACATAATTCA CTCTGAGTGT 4560
TTTATGAGAA AGCCTTCTTT TGGGTCAAC AGTTTCTCTA TGCTTTGAAA CAGAAAAATA 4620
TGTACCAAGA ATCTTGTTT GCCTTCCAGA AAACAAAAC GCATTTCACT TTCCCGGTGT 4680
TCCCACTGT ATCTAGGCAA CATAGTATTC ATGACTATGG ATAAACTAAA CACGTGACAC 4740
AAACACACAC AAAAGGGAAC CCAGCTCTAA TACATTCCAA CTGATATAGC ATGCATCTGT 4800
TTATTCTATA GTTATTAAGT TCTTTAAAT GTAAAGCCAT GCTGGAAAT AATACTGCTG 4860
AGATACATAC AGAATTACTG TAACTGATTA CACTTGGTAA TTGTACTAAA GCCAAACATA 4920
TATATACTAT TAAAAAGGTT TACAGAATTT TATGGTGCAT TACGTGGGCA TTGCTTTTT 4980
20 AGATGCCCAA ATCCTTAGAT CTGGCATGTT AGCCCTTCCT CCAATTATAA GAGGATATGA 5040
ACCAAAAAA AAAAAAAA AA

Seq ID NO: 427 Protein sequence
Protein Accession #: NP_003465

25 1 11 21 31 41 51
MAARPLVPSP ARALLLALAG ALLAPCEARG VSLMNEGRAD EVVSASVRSQ DLWIPVKSFD 60
SKNHPEVLNI RLQRESKELI INLERNEGLI ASSFTETHYL QDGTDVSLAR NYTVILGHCI 120
30 YHGHVRGYSD SAVSLSTCSG LRGLIVFENE SYVLEPMKSA TNRYKLFPAK KLSVRSQSG 180
SHHNTPLNLA KNVFPSPSQ WARRHKRETL KATKYVELVI VADNREFQRQ GKDLKVKQR 240
LIEIANHVOK FYRPLNIRIV LVGVEVNDM DKCSVSQDPF TSLHEFLDWR KMKLLPRKSH 300
DNAQLVSGVY FGTTTIGMAP IMSCTADQS GGIVMDHSDN PLGAAVTLAH ELGHNFGMNH 360
35 DTLDRGCSQ MAVESKGCIM NASTGYPPPM VFSSCSRKDL ETSLEKMGV CLFNLPVRE 420
SFGGQKCGNR FVEGEEDCD GEPEECMNR CNAATCTLKP DAVCAHGLCC EDCQLKPAQT 480
ACRDSNSNCD LPEFCTGASP HCPANVYLHD GHSQDQVDGY CYNGICQTHE QQCVTLMGPG 540
AKPAPGICFE RVNSAGDPYG NCGKVSXSSF AKCEMRDAKC GKIQCGGAS RPIVGTNAV 600
IETNIPLOQG GILRCGTHV YLGDDMPDPG LVLAGTKCAD GKICLNRCQ NISVPGVHEC 660
40 AMQCHGRGVC NNRNCHCEA HWAPFPCKDF GFGSSTDSGP IRQADNQLT IGLVLTILCL 720
LAAGFVVYLL RKTLLIRLLFT NKKTITIEKLR CVRPSRPPRG PQPCQHLGH LGKGLMRKPP 780
DSYFPKDNPR RLLQCNQNDI SRPLNGLNVP QPQSTQVLP PLHRAFRAPS VPARPLPAKP 840
ALRQAQGTCK PNPPQKPLPA DPLARTTRLT HALARTPGQW ETGLRLAPLR PAPQYPHQVP 900
RSTHTAYIK

45 Seq ID NO: 428 DNA sequence
Nucleic Acid Accession #: NM_003714
Coding sequence: 135..1043

50 1 11 21 31 41 51
GAGGAGGAGG GAAAGGCGA GCAAAAAGGA AGAGTGGGAG GAGGAGGGGA AGCGGCGAAG 60
GAGGAAGAGG AGGAGGAGGA AGAGGGGAGC ACAAAGGATC CAGGTCTCCC GACCGGAGGT 120
TAATACCAAG AACCATGTGT GCGAGCGGCG TGGGCCAGTT CATGACCCCTG GCTTTGGTGT 180
55 TGGCCACCTT TGACCGCGCG CGGGGAGCCG ACGCCACCAA CCCACCGAG GGTCCCCAAG 240
ACAGGAGCTC CCAGCAGAAA GCGCGCTGT CCTGCAGAA TACAGCGAG ATCCAGCACT 300
GTTTGGTCAA CGCTGGCGAT GTGGGGTGTG GCGTGTGTA ATGTTTCGAG AACAACTCTT 360
GTGAGATTCC GGGCTTACAT GGGATTGCA TGACTTTTCT GCACAACGCT GGAATAATTG 420
ATGCCAGGG CAAGTCATT ATCAAAGACG CCTTGAAATG TAAGGCCACG GCTCTGGGCG 480
ACAGGTTCCG CTGATCAAGC CGGAAGTGCC CGGCCATCAG GGAATAGTG TCCAGTTGTC 540
60 AGCGGGAATG CTACCTCAAG CACGACCTGT GCGCGGCTGC CCAGGAGAAC ACCCGGTTGA 600
TAGTGGAGAT GATCCATTTC AAGGACTTGC TGCTGCACGA ACCCTACGTG GACCTCGTGA 660
ACTTGCTGCT GACCTGTGGG GAGGAGGTGA AGGAGGCCAT CACCCACAGC GTGCAGGTTT 720
AGTGTGAGCA GAACCTGGGA AGCCTGTGCT CCATCTGAG CTTCTGACCC TCGGCCATCC 780
AGAAGCCTCC CACGCGCCCC CCCGAGCGCC AGCCCCAGGT GGACAGAAC AAGCTCTCCA 840
65 GGGCCACCA CGGGAAGCA GGACATCACC TCCAGAGCC CAGCAGTAGG GAGACTGGCC 900
GAGGTGCCAA GGTGAGCGA GGTAGCAAGA GCCACCCAAA CGCCATGCC CGAGGCAGAG 960
TGGGGGCGCT TGGGGCTCAG GGACCTTCCG GAAGCAGCGA GTGGGAAGAC GAACAGTCTG 1020
AGTATTCTGA TATCCGAGG TGAATGAAA GGCCTGGCCA CGAAATCTTT CCTCCACGCC 1080
70 GTCCATTTC TTATCTATGG ACATTCCAAA ACATTACCA TTAGAGAGGG GGGATGTAC 1140
ACGAGAGATC CTGTGGGGA TGTGACTTC ATCGAGGTGT GTGTTCCGG AACGGACAGG 1200
TGAGATGGAG ACCCTGGGG CCGTGGGTC TCAGGGGTGC CTGTTGAATT CTGACTTAC 1260
ACGTACTCAA GGGAGCGCG CCGGTTATC CTGTAACCTT TGTCTCTTT CCATCTGTG 1320
AGTCAGTGGG TGTGGGCGC TCTGTTGTG GGGAGGTGAA CCAGGGAGGG GCAGGGCAG 1380
GCAGGCGCCC CAGAGCTGGG CCACACAGTG GGTGCTGGG CTCGCCCCGA AGCTTCTGTT 1440
75 CGAGCAGCCT CTGTGCTGT CTCCGCGAA GTCAGGGCGG CTGATTCCA GACAGGAGT 1500
GAATGTAAAA ATAAATATCG CTTAGAAATG AGGAGAAGGG TGGAGAGGAG GCAGGGGCGG 1560
AGGGGGTGTG TGGTGCCAAA CTGAAATTC GTTCTTGTG TGGGGCCTTG CGGTTCCAG 1620
CTCTTGCGGA GGGTGGAGGG AGGAGTGTCA TTTCTATGT TAATTCTGA GCCATGTAC 1680
TGTCTGGGCT GGGGGGGA CA CTGTCCAAG GAGTGGCCCC TATGAGTTTA TATTTTAAC 1740
80 ACTGCTTCAA ATCTCGATT CACTTTTTT ATTTATCCAG TTATATCTAC ATATCTGTCA 1800
TCTAAATAAA TGCTTTCAA ACAAGCAAC TGGGTCAATA AAACCACTC AAAGGGGTT 1860
TAAAAAAA AAAACAGCC CATCTTTGA GGTGATTTT TCTTTTTTT AAGTTCTATT 1920
TAAAAAGTA TCAACAGCG ACATAGCCAT ACATCTGACT GCCTGACATG GACTCCTGCC 1980
CACTTGGGGG AAACCTTATA CCCAGAGGAA AATACACACC TGGGGAGTAC ATTTGACAAA 2040
85 TTTCCCTTAG GATTTCTGTTA TCTCACCTTG ACCCTCAGCC AAGATTGGTA AAGCTGCTC 2100
CTGGCGATTC CAGGAGACCC AGCTGGAAC CTGGCTTCTC CATGTGAGGG GATGGGAAAG 2160
GAAAGAAGAG AATGAAGACT ACTTAGTAAT TCCATCAGG AAATGCTGAC CTTTACATA 2220

AAATCAAGGA GACTGCTGAA AATCTCTAAG GGACAGGATT TTCCAGATCC TAATTGGAAA 2280
TTTAGCAATA AGGAGAGGAG TCCAAGGGGA CAAATAAAGG CAGAGAGAGA GAGAGAGAGA 2340
GGGAGAGGAA GAAAAGAGAG AGAGAAAAGA GCCTCGTGCC

5 Seq ID NO: 429 Protein sequence
Protein Accession #: NP_003705

10 1 11 21 31 41 51
MCAERLGQFM TLALVLATFD PARGTDATNP PEGPQDRSSQ QXGRLSLQNT AEIQHCLVNA 60
GDVGCQVFBC FENNSCEIRG LHGICMTFLH NAGKFDAGGK SPIKDALKCK AHALHRHFGC 120
ISRRCPAIRE MVSQIQRECY LKHDLCAAAQ ENTRVIVEMI HFKDLLLHEP YVDLVNLLLT 180
CGEEVKCAIT HSVQVQCEQN WGSLSLSILSF CTSAIQKPPT APPERQPVQD RTKLSRAHHG 240
EAGEHLPEPS SRETGRGAKG ERGSKSHFNA HARGRVGGLG AQQPSGGSSEW EDEQSEYSDI 300
RR

Seq ID NO: 430 DNA sequence
Nucleic Acid Accession #: NM_005940
Coding sequence: 23..1489

20 1 11 21 31 41 51
AAGCCAGCA GCCCGGGGCG GATGGCTCC GCGCGCTCG CTCCGAGCG CGGCCGCGCG 60
CGCCCTCTCT GCGCCGATGC TGCTGCTGCT GCTCCAGCG CCGCGCTGC TGGCCCGGGC 120
25 TCTGCCGCG GAGCTCCACC ACCTCCATGC CGAGAGGAG GGGCCACAGC CCTGGCATGC 180
AGCCCTGCCG AGTAGCCCGG CACCTGCCCG TGCCACGCG GAAGCCCCCG GGCTCGCCAG 240
CAGCCTCAGG CCTCCCGCTG GTGGCGTGCC CGACCCATCT GATGGGCTGA GTGCCCGCAA 300
CCGACAGAAG AGGTTGCTGC TTTCTGGCGG GCGCTGGGAG AAGACGGACC TCACCTACAG 360
GATCCTTCGG TCCCATGGG AGTTGGTGCA GGAGCAGGTG CCGCAGACGA TGGCAGAGGC 420
30 CCTAAAGGTA TGGAGCGATG TGAGGCCACT CACCTTTACT GAGGTGCACG AGGGCCGTGC 480
TGACATCATG ATCGACTTCG CCAGGTACTG GCATGGGAC GACCTGCCGT TTGATGGGCC 540
TGGGGGCATC CTGGCCCATG CCTTCTTCCC CAAGACTCAC CGAGAAGGGG ATGTCCACTT 600
CGACTATGAT GAGACCTTGA CTATCGGGGA TGACAGGGGC ACAGACCTGC TGCAGGTGGC 660
AGCCCATGAA TTTGGCCAGC TGCTGGGGCT GCAGCACACA ACAGCAGCCA AGGCCCTGAT 720
35 GTCCGCTCTC TACACTTTC GCTACCCACT GAGTCTCAGC CCAGATGACT GCAGGGGGGT 780
TCAACACCTA TATGGCCAGC CCTGGCCAC TGTCACCTCC AGGACCCAGC CCCTGGGGCC 840
CCAGGCTGGG ATAGACACCA ATGAGATTGC ACOGCTGGAG CCAGACGCCC CGCCAGATGC 900
CTGTGAGGCC TCCTTTGAGC CGGTCTCCAC CATCCGAGGC GAGCTCTTTT TCTTCAAGC 960
40 GGGCTTTTGG TGGGCGCTCC GTGGGGGCCA GCTGCAGCCC GGCTACCCAG CATTTGCCCTC 1020
TCGCCACTGG CAGCGACTGC CCAGCCCTGT GGAGCCTGCC TTGAGGATG CCCAGGGCCA 1080
CATTGGTTTC TTCCAAAGTG CTCAGTACTG GGTGTACGAC GGTGAAAAGC CAGTCTCTGG 1140
CCCGCAGCCC CTCACGAGC TGGGCGCTGT GAGGTTCGCG GTCCATGCTG CCTTGGTCTG 1200
GGGTCCCGAG AAGAACAGA TCTACTCTCT CCGAGGCAGG GACTACTGGC GTTTCACACC 1260
45 CAGCACCCGG CGTGTAGACA GTCCCGTGCC CCGCAGGGCC ACTGACTGGA GAGGGGTGCC 1320
CTCTGAGATC GAGCTGCTCT TCCAGGATGC TGATGGCTAT GCCTACTTCC TGGCGGGCCG 1380
CCTCTACTGG AAGTTTGACC CTGTGAAGGT GAAGGCTCTG GAAGGCTTCC CCCGTCTCGT 1440
GGGTCTGAC TTCTTTGGCT GTGCCGAGCC TGCCCAACT TCTCTCTGAC CATGGCTTGG 1500
ATGCCCTCAG GGTGTCTGAC CCCTGCCAGG CCACGAATAT CAGGCTAGAG ACCCATGGCC 1560
50 ATCTTTGTGG CTGTGGGCAC CAGGCATGGG ACTGAGCCCA TGTCTCTCTG AGGGGGATGG 1620
GGTGGGGTAC AACCACCATG ACAAATGCCG GGAGGGCCAC GCAGGTCTGT GTCACTTGGC 1680
AGCGACTGTC TCAGACTGGG CAGGGAGGCT TTGGCATGAC TTAAGAGGAA GGGCAGTCTT 1740
GGGACCGCTG ATGCAGGTCC TGGCAACCT GCGTCCCTG TCTCATCCCT GTCCCTCAGG 1800
GTAGCACCAT GGCAGGACTG GGGGAACCTG AGTGTCTCTG CTGTATCCCT GTTGTGAGGT 1860
55 TCCTTCAGAG GGTCTGCACT GAAGCAAGGG TGCTGGGGCC CCATGGGCCTT CAGCCCTGGC 1920
TGAGCAACTG GGTCTGAGGG CAGGGCCACT TCCTGAGGTC AGGTCTTGGT AGGTGCTCTG 1980
ATCTGTCTGC CTTCTGGCTG ACAATCTCTG AAATCTGTTC TCCAGAAATC AGGCCAAAAA 2040
GTTACAGLFF FKAGFVWRLR GGQLQPGYPA LASRHWQGLP SFVDAAFEDA QGHIWFPQGA 2100
CAACATACCT CAATCTGTGC CCAGGCCGGA TCCTCTGAA GCCCTTTTTC CAGCACTGCT 2160
60 ATCTCTCAAA GCCATGTGAA ATGTGTGTAC AGTGTGTATA AACCTTCTTC TTCTTTTTTT 2220
TTTTTAAACT GAGGATGTGC ATTAACACA GTTGTTTTCT

Seq ID NO: 431 Protein sequence
Protein Accession #: NP_005931

65 1 11 21 31 41 51
MAPAAWLRS AARALLPPML LLLLQPPPLL ARALPPDVHH LHAERRGPQ WHAALPSSPA 60
PAPATQZAPR PASSLRPPRC GVPDPSDGLS ARNRQKRFVL SGRWKEKDL TYRILRFPWQ 120
70 LVQSQVRQTM AEALKVNSDV TPLTFTEVHE GRADIMIDFA RYWHGDDLFP DPGGILAH 180
FFPKTHREGD VHPDYDETWT IGDDQGTDL LQVAHEFGHV LGLQHTTAAK ALMSAFYTPR 240
YPLSLSPDDC RGVOHLYGQP WPTVTSRTPA LGPOAGIDTN EIAPLEPDAP PDACEASFDA 300
VSTIRGELEF FKAGFVWRLR GGQLQPGYPA LASRHWQGLP SFVDAAFEDA QGHIWFPQGA 360
QYVWYDGEKP VLGPAPLTEL GLVRFPVHAA LVHGPENKNI YFFRGRDYWR FHPSTRRVDS 420
75 PVPRRATDWR GVPSEIDAAF QDADGAYFL RGRLYWKFPD VKVKALEGFP RLVGPDFPFGC 480
AEPANTFL

Seq ID NO: 432 DNA sequence
Nucleic Acid Accession #: NM_024022
Coding sequence: 202..1563

80 1 11 21 31 41 51
ACCGGGCACC GGAAGGCTCG GGTACTTTCG TTCTTAATTA GGTATGCCCC GTGTGAGCCA 60
GGAAGGGCTG GTGTTTATGG GAAGCCAGTA ACACCTGTGC CTACTATCTC TTCCGTGGTG 120
85 CCATCTACAT TTTTGGGACT CGGGAATTAT GAGGTAGAGG TGGAGGCGGA GCCGATGTC 180
AGAGTCTCTG AATAGTTCAC CATGGGGGAA AATGATCCGC CTGCTGTGTA AGCCCCCTTC 240
TCATTCGATG CGCTTTTGGG CCTTGATGAT TTGAAAATAA GTCCTGTGTC ACCAGATGCA 300

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

GATGCTGTTG CTGCACAGAT CCTGTCACTG CTGCCATTGA AGTTTTTTTC AATCATCGTC 360
ATTGGGATCA TTGCATTGAT ATTAGCACTG GCCATTGGTC TGGGCATCCA CTTCGACTGC 420
TCAGGGAATG ACAGATGTGC CTCATCCTTT AAGTGTATCG AGCTGATAGC TCGATGTGAC 480
GGAGTCTCGG ATTGCAAGA CGGGGAGGAC GAGTACCGCT GTGTCCGGGT GGGTGGTCAG 540
AATGCCGTGC TCCAGGTGTT CACAGCTGCT TCGTGGGAAGA CCATGTGCTC CGATGACTGG 600
AAGGGTCACT ACGCAATGT TGCCTGTGCC CAACCTGGGT TCCCAAGCTA TGTGAGTTCA 660
GATAACCTCA GAGTGAAGTC GCTGGAGGGG CAGTTCCGGG AGGAGTTTGT GTCCATCGAT 720
CACTCTTGCC CAGATGACAA GGTGACTGCA TTACACCACT CAGTATATGT GAGGGAGGGA 780
TGTGCTCTCG GCCACGTGGT TACCTTGCAG TGCACAGCCT GTGGTCATAG AAGGGGCTAC 840
AGCTCAAGCA TCGTGGGTGG AAACATGTCC TTGCTCTCGC AGTGGCCCTG GCAGGCCAGC 900
CTTCAGTTCC AGGCTACCA CTTGTGCGGG GGCTCTGTCA TCAAGCCCTT GTGGATCATC 960
ACTGCTGCAC ACTGTGTTTA TGAATTGTAC CTCCCAAGT CATGACCAT CCAGGTGGGT 1020
CTAGTTTCCC TGTGTGACAA TCCAGCCCCA TCCCACTGG TGGAGAAGAT TGTCTACCAC 1080
AGCAAGTCA AGCCAAAGAG GCTGGGCAAT GACATCGCCC TTATGAAGCT GGCCGGGCCA 1140
CTCACGTTCA ATGAATGAT CCAGCCTGTG TGCCTGCCCA ACTCTGAAGA GAACTTCCCC 1200
GATGGAAATG TGTGCTGGAC GTCAGGATGG GGGGCCACAG AGGATGGAGG TGAAGCCTCC 1260
CCTGTCTGTA ACCAAGCGGC CGTCCCTTTG ATTTCCAACA AGATCTGCAA CCACAGGGAC 1320
GTGTACGGTG GCATCATCTC CCCCTCCATG CTCTCGCGGG GCTACCTGAC GGGTGGCGTG 1380
GACAGCTGCC AGCGGGACAG CGGGGGGCCC CTGGTGTGTC AAGAGAGGAG GCTGTGGAAG 1440
TTAGTGGGAG CGACAGCTT TGGCATCGGC TGGCAGAGG TGAACAAGCC TGGGGTGTAC 1500
ACCCGTGTCA CCTCTTCTCT GGAAGTGGAT CACGAGCAGA TGGAGAGAGA CTTAAAAACC 1560
TGAAGAGGAA GGGGACAAAT AGCCACCTGA GTTCTCTGAG TGAAGAAGAC AGCCCGATCC 1620
TCCCCTGGAC TCCCGTGTAG GAACCTGCAC ACGAGCAGAC ACCCTTGGAG CTCTGAGTTC 1680
CGGCACCACT AGCAGGCCCG AAAGAGGCAC CCTTCCATCT GATTCCAGCA CAACCTTCAA 1740
GCTGCTTTT GTTTTTGTGT TTTTGTAGGT GGAGTCTCGC TCTGTGCCC AGGCTGGAGT 1800
GCAGTGGGGA AATCCCTGCT CACTGCAGCC TCCGCTTCCC TGGTTCAAGC GATTCTCTTG 1860
CCTCAGCTTC CCAAGTAGCT GGGACACAG GTGCCCGCCA CCACACCCAA CTAATTTTGT 1920
TATTTTTAGT AGAGACAGGG TTTCAACATG TTGGCCAGGC TGCTCTCAA CCCCTGACCT 1980
CAATGATGT GCCTGCTTCA GCCTCCCACT GTGCTGGGAT TACAGGCATG GGCCACCAAG 2040
CCTAGCCTCA CGCTCCTTTC TGATCTTCAC TAAGAACAAA AGAAGCAGCA ACTTGCAAGG 2100
GCGGCTTTTC CCACTGGTCC ATCTGGTTTT CTCTCCAGGG GTCTTGCAA ATTCCTGAGC 2160
AGATAAGCAG TTATGTGACC TCACGTGCAA AGCCACCAAC AGCCACTAG AAAAGAGGCA 2220
CCAGCCCAAG AGTGCAGAAC TGCACTCACT GCACGTTTT CTCTTAGGG ACCAGAACCA 2280
AACCACCTCT TTACTTTC AAGACTTATT TTCACATGTG GGGAGGTTAA TCTAGGAATG 2340
ACTCGTTTAA GGCTATTITT CATGATTCTT TTGTAGCATT TGGTGTGTA CGTATTATTG 2400
TCCTTTGATT CCAATAATAA TGTTCCTTC CCTCAAAAAA AAAAAA AAAA 2460
AAAA

Seq ID NO: 433 Protein sequence
Protein Accession #: NP_076927

1 11 21 31 41 51
MGENDPPAVE APFSFRSLPG LDDLKISFVA PDADAVAQAI LSLPLKFFP ITVIGIILI 60
45 LALAIGLGIH FDCSGKYRRC SSFKIELIA RCDGVSDCKD GEDEYRCVRV GQNAVLQVP 120
TAASWTKMCS DDWKGHYANV ACAQLGFPSY VSSDNLRVSS LEGQFREEFV SIDHLLPDDK 180
VTALHHSVTV REGCASHVUV TLOCTACGHR RGYSSRIVGG NMSLLSQWPW QASLQFQGYH 240
LCGGSVITPL WIITAHRCVY DLYLFPKSWTI QVGLVSLLDN PAPSHLVEKI VYHSKYKPKR 300
50 LGNDIALMLK ASPITFNEMI QPVCLPNSEB NFPGDKVCWT SGWGATEDGG DASPVNLBA 360
VPLISNKICN HRDVGYYIIS PSMCLAGYLT GGVDSQCGDS GGPLVCQERR LKLVGATSP 420
GIGCAEVNKP GVTYTRVTSPL DWIHEQMERD LKT

Seq ID NO: 434 DNA sequence
Nucleic Acid Accession #: NM_000493.2
Coding sequence: 97..2139

1 11 21 31 41 51
60 CACCTTCTGC ACTGCTCATC TGGGCAGAGG AAGCTTCAGA AAGCTGCCAA GGCACCATCT 60
CCAGGAACCTC CCAGCAGCGA GAATCCATCT GAGAATATGC TGCCACAAAT ACCCTTTTGT 120
CTGCTAGTAT CCTTGAACCT GGTTCATGGA GTGTTTTACG CTGAACGATA CCAAAATGCC 180
ACAGGCATAA AAGGCCCACT ACCCAACACC AAGACACAGT TCTTCATTCC CTACACCATA 240
AAGAGTAAAG GTATAGCAGT AAGAGGAGAG CAAGGTACTC CTGTTCCACC AGGCCCTGCT 300
65 GGACCTCGAG GGCACCCAGG TCCTTCTGGA CCACCAGGAA AACCAAGCTA CGCAAGTCTCT 360
GGACTCCAAAG GAGAGCCAGG GTTGCCAGGA CCACCGGAC CATCAGCTGT AGGGAAACCA 420
GGTGTGCCAG GACTCCACAG AAAACCAAGG GAGAGAGGAC CATATGGACC AAAAGGAGAT 480
GTTGGACCAAG CTGGCCTACC AGGACCCCGG GGCCCAACAG GACCACCTGG AATCCCTGGA 540
CCGGCTGGAA TTTCTGTGCC AGGAAACCTT GGACAAACAG GACCCACAGG AGCCCCAGGA 600
70 CCCAGGGGCT TTCTTGGAGA AAAGGGTGCA CCAGGAGTCC CTGGTATGAA TGGACAGAAA 660
GGGGAAATGG GATATGGTGC TCCTGGTGGT CCAGGTGAGA GGGGTCTTCC AGGCCCTCAG 720
GGTCCACAG GACCATCTGG CCCTCCTGGA GTGGGAAAAA GAGGTGAAAA TGGGGTTCCA 780
GGACAGCCAG GCATCAAAGG TGATAGAGGT TTTCCGGGAG AATGGGACC AATTGGCCCA 840
CCAGGTCCCC AAGGCCCTCC TGGGAAACGA GGGCCAGAG GCATTGGAAA GCCAGGAGCT 900
75 GCTGGAGCCC CAGGCCAGCC AGGGATTCCA GGAACAAAAG GTCTCCCTGG GGCTCCAGGA 960
ATAGCTGGAC CCCAGGGGCC TCCTGGCTTT GGGAAACCAAG GCTTGCAGG CTGAAGGGA 1020
GAAAGAGGAC CTGCTGGCCT TCCTGGGGGT CCAGGTGCCA AAGGGGAACA AGGGCCAGCA 1080
GGTCTTCTCG GGAAGCCAGG TCTGACTGGA CCCCCTGGGA ATATGGGACC CCAAGGACCA 1140
AAAGGCATCC CGGTAGCCA TGGTCTCCCA GGCCCTAAAG GTGAGACAGG GCCAGCTGGG 1200
CCTGCAGGAT ACCCTGGGGC TAAGGGTGAA AGGGGTTCCC CTGGGTGAGA TGGAAAAACA 1260
80 GGGTACCCAG GAAACACAGG TCTCGATGGT CCTAAGGGTA ACCCAGGGTT ACCAGGTCCA 1320
AAAGGTGATC CTGAGTTGG AGGACCTCCT GGTCTCCAG GCCTGTGGG CCCAGCAGGA 1380
GCAAGGGAA TCCCGGCA CAATGGAGAG GCTGGCCCAA GAGGTGCCCC TGGAAATACA 1440
GGTACTAGAG GCCCTATTGG GCCACCAGSC ATTCCAGGAT TCCTGGGTC TAAAGGGGAT 1500
CCAGGAAGTC CGGTCTCTCC TGGCCAGGCT GGCATAGCAA CTAAGGGCCT CAATGGACCC 1560
85 ACCGGGCCAC CAGGGCCCTC AGGTCCAAGA GGCCACTCTG GAGAGCCTGG TCTTCCAGGG 1620
CCCCCTGGGC CTCCAGGCC ACCAGGTCAA GCAATCATGC CTGAGGGTTT TATAAGGCA 1680
GGCCAAAGGC CCACTCTTTC TGGGACCCCT CTGTGATGTC CCAACAGGG GGTAAACAGGA 1740

	ATGCTGTGT	CTGCTTTTAC	TGTTATTCTC	TCCAAAGCTT	ACCCAGCAAT	AGGAACCTCC	1800
	ATACCATTTG	ATAAATTTT	GTATAACAGG	CAACAGCATT	ATGACCCAAG	GACTGGAATC	1860
	TTTACTTGTC	AGATACCAGG	AATATACTAT	TTTTCATACC	ACGTGCATGT	GAAAGGGACT	1920
5	CATGTTTGGG	TAGGCTGTGA	TAAGAATGGC	ACCCCTGTAA	TGTACACCTA	TGATGAATAC	1980
	ACCAAAGGCT	ACCTGGATCA	GGCTTCAGGG	AGTGCCATCA	TCGATCTCAC	AGAAAATGAC	2040
	CAGGTGTGGC	TCCAGCTTCC	CAATGCCGAG	TCAAATGGCC	TATACTCCTC	TGAGTATGTC	2100
	CACCTCTCTT	TCTCAGGATT	CCTAGTGGCT	CCAATGTGAG	TACACCCAC	AGAGCTAATC	2160
	TAAATCTGTG	GCTAGAAAAA	GCATTCTCTA	ACTCTACCCC	ACCCATACAA	ATGCATATGG	2220
	AGGTAGGCTG	AAAAGAATGT	AAITTTTATT	TTCTGAAATA	CAGATTTGAG	CTATCAGACC	2280
10	AACAAACCTT	CCCCCTGAAA	AGTGAGCAGC	AACGTAAAAA	CGTATGTGAA	GCCTCTCTTG	2340
	AATTTCTAGT	TAGCAATCTT	AAGGCTCTTT	AAGGTTTCT	CCAATATTAA	AAAAATATCAC	2400
	CAAAGAAGTC	CTGCTATGTT	AAAAACAAAC	AACAAAAAAC	AAAGCAACAA	AAAAAAAAT	2460
	TAAAAAATAA	AACAGAAATA	GAGCTCTAAG	TTATGTGAAA	TTTGATTGGA	GAAACTCGGC	2520
	ATTTCTTTT	TAAAAAGCC	TGTTTCTAAC	TATGAATATG	AGAACTTCTA	GGAAACATCC	2580
15	AGGAGGTATC	ATATAACTTT	GTAGAACCTA	AATACTTGAA	TATTCAAATT	TAAAAGACAC	2640
	TGTATCCCCC	AAAATATTTT	TGATGGTGCA	CTACTCTGAG	GCCTGTATGG	CCCCTTTCAT	2700
	CAATATCTAT	TCAAATATAC	AGGTGCATAT	ATACTTGTTA	AAGCTCTTAT	ATAAAAAAGC	2760
	CCCAAAATAT	TGAAGTTTAT	CTGAAATGCA	AGGTGCTTTC	ATCAATGAAC	CTTTTCAAAA	2820
	CTTTTCTATG	ATTGCAGAGA	AGCTTTTAT	ATACCCAGCA	TAACTTGGAA	ACAGGTATCT	2880
20	GACCTATTCT	TATTTAGTTA	ACACAAGTGT	GATTAATTTG	ATTTCTTTAA	TTCTTATTG	2940
	AATCTTATGT	GATATGATT	TCTGATTTA	CAGAACATTA	GCACATGTAC	CTTGTGCTC	3000
	CCATTCAAGT	GAAGTATATA	TTTACACTGA	GGGTTTCAAA	ATTCGACTAG	AAGTGGAGAT	3060
	ATATTATTTA	TTTATGCAC	GTACTGTATT	TTTATATTGC	TGTTTAAAC	TTTTAAGCTG	3120
	TGCTCACTT	ATTAAGCACC	AAAATGTTTT	ACCTACTCCT	TATTTACGAC	ACAATAAAAT	3180
25	AACATCAATA	GATTTTATAG	CTGAATTAAT	TTGAAAGCAG	CAATTGCTG	TTCTCAACCA	3240
	TTCTTTCAAG	GCTTTTCATT	CGACACAATA	AAATAACATC	AATAG		

Seq ID NO: 435 Protein sequence
Protein Accession #: NP_000484.2

	1	11	21	31	41	51	
	MLPQIPFLLL	VSLNLVHGVS	YAERYQMPTG	IKGSLPNTKT	QFFIFPTIKS	KGIIVRGEQG	60
	TPGPPGAPG	RGHGPGSGPP	GKPGYGSPLG	QGEPLGPGPP	GPSAVGKPGV	PGLPGKPGER	120
35	GPYGPKEGVS	PAGLPGPRGP	PGPPGIPGPA	GISVPGKPGQ	QGPTGAPGER	GPPGKPGAPG	180
	VPMNGMGKGE	MGYGAAPRGP	ERGLPGPQGP	TGSPGPPGVG	KRGNGVPGQ	PGIKGDRGFP	240
	GEMGPIGPPG	PQPPGGERGP	EGIGKPGAAG	APGQPGIPGT	KGLPGAPGIA	GPPGPPGPGK	300
	PGLPGLKGER	GPAGLPGGGP	AKGEQGPAGL	PGKPLGTGPP	GNGPGQPGK	IPGSHGLPGP	360
	KGETGPAGPA	GYPGAAGERG	SPGSDGKPGY	PGKPLGDGPK	GNGPLPGPKG	DPGVGGPPGL	420
40	PGPVGPAGAK	GMPGHNGEAG	PRGAPGIPGT	RGPFGPPGIP	GFPGSKGDPG	SPGPPGPGAGI	480
	ATKGLNGPTG	PPGPPGPRGH	SGEPGLPGPP	GPPGPPGQAV	MPEGFIKAGQ	RPSLSGTPLV	540
	SANQGVGTMP	VSAFTVILSK	AYPAIGTPIP	FDKILYNRQQ	HYDPRGTGFT	CQIPGIYFYS	600
	YEVGVKGVTH	WVGLYKNGTP	VMYTYDEYTK	GYLDQASGSA	IIDLTEVDQV	WLQLPNAESN	660
	GLYSSEKVS	SFGSFLVAPM					

Seq ID NO: 436 DNA sequence
Nucleic Acid Accession #: XM_062811
Coding sequence: 1..888

	1	11	21	31	41	51	
	ATGTGGGGCG	CTGCGCGCTC	GTCCGTCTCC	TCATCCTGGA	ACGCCGCTTC	GCTCCTGCAG	60
	CTGCTGCTGG	CTGCGCTGCT	GGCGGCGGGG	GCGAGGGCCA	GCGGCGAGTA	CTGCCACGGC	120
	TGGCTGACCG	CGCAGGCGGT	CTGGCGCATC	GGCTTCCAGT	GTCCCGAGCG	CTTCGACGGC	180
55	GGCGACGCCA	CCATCTGCTG	CGGCAGCTGC	CGCTTGCGCT	ACTGCTGCTC	CAGCGCCGAG	240
	GCGCGCTCG	ACCAAGGCGG	CTGCGACAAT	GACCGCCAGC	AGGGCGCTGG	CGAGCCTGGC	300
	CGGGCGGACA	AAGACGCGCC	CGACGCTCG	GCACTGCCCA	TCTACGTGCC	GTCTCTCAT	360
	GTGTGCTCG	TGTTTGTGCG	CTTTATCATC	TTGGGTCCCC	TGGTGGCAGC	CTGTGTCTGC	420
	AGATGTCTCC	GGCTTAAGCA	GGATCCCCAG	CAGAGCCGAG	CCCCAGGGGG	TAACCGCTTG	480
60	ATGGAGACCA	TCCCATGAT	CCCCAGTGCC	AGCACTCCCC	GGGGTCTGTC	CTCACGCCAG	540
	TCCAGCACCA	CTGCCAGTTC	CAGCTCCAGC	GCCAACTCAG	GGGCCCGGGC	GCCCCAACAA	600
	AGGTACACAG	CCAATCTGTT	CTTGCCGGAA	GGGACCATGA	ACAACGTGTA	TGTCAACATG	660
	CCCACGAATT	TCTCTGTGCT	GAACTGTGAG	CAGGCCACCC	AGATTGTGCC	ACATCAAGGG	720
	CAGTATCTGC	ATCCCCCATA	CGTGGGTGAC	ACGCTGTCAG	ACGACTCTGT	GCCCATGACA	780
65	GCTGTGCCAC	CTTTCATGGA	CGGCTGTCAG	CCTGGCTACA	GGCAGATTCA	GTCCCCCTTC	840
	CCTCACACCA	ACAGTGAACA	GAAGATGTAC	CCAGCGGTGA	CTGTATAA		

Seq ID NO: 437 Protein sequence
Protein Accession #: XP_062811

	1	11	21	31	41	51	
	MWGARRSVS	SSWNAASLLQ	LLLAALLAAG	ARASGEYCHG	WLDAGQVWRI	GFQCPERFDG	60
	GDATICCGSC	ALRYCCSSAE	ARLDQGGCDN	DRQQGAGEPG	RADKDGPOGS	AVPIYVPLI	120
75	VGSVFVAFII	LGLSLVAACC	RCLRPKQDPQ	QSRAPGNNRL	METIPMIPSA	STSRGSSSRQ	180
	SSTAASSSSS	ANSGARAPPT	RSQTNCCLEP	GTNNVYVNM	PTNFSVLNCO	QATQIVPHQG	240
	QYLHPYVGY	TQVHDSVPMT	AVPPFMDGLQ	PGYRQIQSPF	PHNSEQKMY	PAVTV	

Seq ID NO: 438 DNA sequence
Nucleic Acid Accession #: NM_004004.1
Coding sequence: 1..681

	1	11	21	31	41	51	
85	ATGGATTGGG	GCACGCTGCA	GACGATCCTG	GGGGGTGTGA	ACAAACACTC	CACCAGCATT	60
	GGAAAGATCT	GGCTACACGT	CCTCTTCATT	TTTCGCAITTA	TGATCTCGT	TGTGGCTGCA	120
	AAGGAGGTGT	GGGGAGATGA	GCAGGCCGAC	TTTGTCTGCA	ACACCTGCA	GCCAGGCTGC	180

5
10
15
20
25
30
35
40
45
50
55
60
65
70

```
AAGAACGTGT GCTACGATCA CTACTTCCCC ATCTCCACCA TCCGGCTATG GGCCTTCGAG 240
CTGATCTTTCG TGTCCAGCCC AGCGCTCCTA GTGGCCATGC ACGTGGCCTA CCGGAGACAT 300
GAGAAAGAGA GGAAGTTCAT CAAGGGGGAG ATAAAGAGTG AATTTAAGGA CATCGAGGAG 360
ATCAAAACCC AGAAGGTCCG CATCGAAGGC TCCTGTGGT GGACCTACAC AAGCAGCATC 420
TTCTTCGGGG TCATCTTCGA AGCGCCCTTC ATGTACGTCT TCTATGTCAT GTACGACGGC 480
TTCTCCATGC AGCGGTCTGT GAAGTGCAAC GCCTGGCCTT GTCCCAACAC TGTGGACTGC 540
TTTGTGTCCC GGCCTCAGGA GAAGACTGTC TTCACAGTGT TCATGATTGC AGTGTCTGGA 600
ATTTCATCC TGCTGAATGT CACTGAATTG TGTATTATGC TAATTAGATA TTGTTCTGGG 660
AAGTCAAAA AGCCAGTTTA A
```

Seq ID NO: 439 Protein sequence
Protein Accession #: NP_003995.1

1 11 21 31 41 51
15
20
25
30
35
40
45
50
55
60
65
70

```
MDWGLTLQIL GGVNKHSTSI GKIWLTVLPI PRIMILVVAA KEVWGDEQAD FVNTLQPGC 60
KNVCYDHYFP ISHRLWALQ LIFVSSPALL VAMHVAYRRH EKRKFIKGE IKSEFKDIEE 120
IKTKQVRIEG SLWNTYTSSI FFRVIFEAAP MYVFYVMYDG FSMQRLVKCN ANPCPNTVDC 180
FVSRPTEKTV FTFVMIIVSG ICILLNVTEL CYLLIRYCSG KSXKPV
```

Seq ID NO: 440 DNA sequence
Nucleic Acid Accession #: XM_061091.1
Coding sequence: 1..2481

1 11 21 31 41 51
25
30
35
40
45
50
55
60
65
70
75
80
85

```
ATGCCAAATA CTTCAGGAAC AACCAGGATT GAAATTTGGC TTCTCCAAGA GCCGCCCGGG 60
CACCGAGGCG TGGTCGCGCG TCTCCTTCCG GTGAGTCCCA GCCCGAGATT GGCTCTGGCG 120
CCCGGGTACC GCGCAGTGCC GCGTCCCGAT GACCGATTCA CGCTCCCGAT GATTGGAGGT 180
CAGATGCATG GTGAGAAGGT AGATCTCTGG AGCCTTGGTG TTCTTTGCTA TGAATTTTAA 240
GTTGGGAAGC CTCCTTTTGA GCGAAACGAA GTCCATGTAA GCAAGAAAC CATCGGGAAG 300
ATTTACAGTG CCAGCAAAAT GATGTGGTGC TCGGCTGCAG TGGACATCAT GTTCTGTGTA 360
GATGGGTCTA ACAGCGTCGG GAAAGGGAGC TTGAAAGGT CCAAGCACTT TGCCATCACA 420
GTCTGTGACG GTCTGGACAT CAGCCCCGAG AGGGTCAGAG TGGGAGCAAT CCAAGTTCAGT 480
TCCACTCCTC ATCTGGAATT CCGCTTGGAT TCATTTTCAA CCAACAGGA AGTGAAGGCA 540
AGAATCAAGA GGATGGTTTT CAAAGGAGGG CGCACGGAGA CGGAACCTGC TCTGAAATAC 600
CTTCTGCACA GAGGGTTGCC TGGAGGCAGA AATGCTTCTG TGCCCCAGAT CCTCATCATC 660
GTCACTGATG GGAAGTCCCA GGGGGATGTG GCACTGCCAT CCAAGCAGCT GAAGGAAAGG 720
GGTGTCACTG TGTTTGCTGT GGGGGTCAGG TTTCCAGGT GGGAGGAGCT GCATGCACATG 780
GCCAGCGAGC CTAGAGGGCA GCACTGCTG TTGGCTGAGC AGGTGGAGGA TGCCACCAAC 840
GGCCTCTTCA GCACCTCTAG CAGCTCGGCC ATCTGCTCCA CGCCCAAGCC AGCTGGGAGC 900
CCGAGGCTTG TCTTCATGGA GGGGTTAATG GGCATCTCTC TGATAGGCCC CTGTGACTCG 960
CAGCCCTGCC AGAATGGAGG CACATGTGTT CCAGAAGGAC TGGACGGCTA CCAAGTGCCTC 1020
TGCCCGCTGG CTTTGGAGG GGAGGCTAAC TGTGCCCTGA AGCTGAGCCT GGAATGCAGG 1080
GTCCAGCTCC TCTTCTGCT GGACAGCTCT GCGGGCACA CTCTGGACGG CTCTCTGCGG 1140
GCCAAAGTCT TCGTGAAGCG GTTGTGCGG GCGGTGCTGA GCGAGGACTC TCGGGCCCGA 1200
GTGGGTGTGT CCACATACAG CAGGGAGCTG CTGTGTGGGG TGCGTGTGGG GGAGTACCAG 1260
GATGTGCTTG ACCTGGTCTG GAGCCTCGAT GGCATTCCCT TCGGTGGTGG CCCCACCTTG 1320
ACGGGCACTG CTTTGGCGCA GCGCGCAGAG CGTGGCTTCG GGAGCCCGAC CAGGACAGGC 1380
CAGGACCGGC CAGCTAGAGT GGTGGTTTTG CTCAGTGAGT CACACTCCGA GGATGAGGTT 1440
GCGGGCCCGC CCGCTCACGC AAGGGCCCGA GAGCTGCTCC TGCTGGGTGT AGGCAGTGAG 1500
GCCGTGCGGG CAGAGCTGGA GGAGATCACA GGCAGCCCAA AGCATGTGAT GGTCTACTCG 1560
GATCCTCAGG ATCTGTTCAA CCAATCCCT GAGCTGCAGG GGAAGCTGTG CAGCCCGGCA 1620
CGGCCAGGCT CCGCGGACCA AGCCCTGGAC CTGCTCTTCA TGTGGGACAC CTCTGCTTCA 1680
GTAGGGCCCG AGAATTTTGC TCAGATGCAG AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT 1740
GAGGTGAACC CTGACGTGAC ACAGGTCCGC CTGTGTGTGT ATGGCAGCCA GGTGCAGACT 1800
GCCTTCGGGC TGGACACCAA ACCACCCGG GCTGTGATGC TGCGGGCCAT TAGCCAGGCC 1860
CCCTACCTAG GTGGGGTGGG CTCAGCCGGC ACCGCCCTCG TGACATCTA TGACAAAGTG 1920
ATGACCGTTC AGAGGGGTGC CCGGCTGCT GTCCCAAGC CTGTGTGTGT GCTCACAGGC 1980
GGGAGAGGCG CAGAGGATGC AGCGGCTCCT GCCCAGAAGC TGAGGAACAA TGGCATCTCT 2040
GTCTTGTGCT TGGGGGTGGG GCCTGTCTTA AGTGAGGGTC TGCGGAGGCT TGCAGGTCCC 2100
CGGGATTCCC TGATCCACGT GGCAGCTTAC GCGACCTGCG GGTACCACCA GGACGTGCTC 2160
ATTGAGTGGC TGTGTGAGA AGCCAAGCAG CCAAGTCAAC TCTGCAAAAC CAGCCCGTGC 2220
ATGAATGAGG GCAGCTGCGT CCTGCAGAA GGGAGCTACC GCTGCAAGTG TCGGATGGC 2280
TGGGAGGGCC CCCACTGCGA GAACCGTGAG TGGAGCTCTT GCTCTGTATG TGTGAGCCAG 2340
AGGATGATTG TTGAGAGGCC CCTGAGGCAC ATGGCTCCCC TGCGAGAGGG CAGCAGCCGT 2400
ACCCCTCCCA GCAACTACAG AGAAGGCTCG GGCAGTAAA TGGTGCTTAC CTCTGGAAT 2460
GTCTGTGCCC CAGTCTCTTA G
```

Seq ID NO: 441 Protein sequence
Protein Accession #: XP_061091.1

1 11 21 31 41 51
75
80
85

```
MPNTSGTTRI EIWLLQEPGP HRLVAALLP VSPSPALALA PGYPPVPAAD DRPTLFMIGG 60
QMHEKVDLW SLGLVLCYFL VGRKPPFEANE VHSKETIGK ISAASKMMC SAAVDIMFL 120
DGSNSVGKGS PERSKHFAIT VCDGLDISPE RVRVGAFQFS STPHLEPLD SPSTQOEKVA 180
RIKRMVFKGG RTETELALKY LLHRLPLPGR NASVPQILII VTDGKSQGDV ALPSKQLKER 240
GUTVFAVGVR FPRWEELHAL ASEPRGQHV LAEQVEDATN GLFSTLSSSA ICSSATPAGS 300
PELVFMERLM GISLIGPCDS QPCQNGGTCV PBGLDGYQCL CPLAFGEAN CALKLSLECR 360
VDLLFLDSS AGTTLDGFLR AKVFKRFVR AVLSEDSRAR VGVATYSREL LVAVPVGEYQ 420
DVPDLVWSLD GIPFRGGPTL TGSALRQAAE RGFSAITRG QDRPRVVVL LTESHSEDEV 480
AGPARHARH ELLLLGVGSE AVRAELEBIT GSPKHMVMS DPQDLFNQIP ELQGLCSRQ 540
RPGCRTQALD LVFMILDSAS VGPNFAQMQ SFVRSCALQF EVNPDVTQVG LVVYGSQVQT 600
APGLDTKPTR AAMLRAISQA PYLGGVGSAG TALLHIYDRV MTVQRGARPG VPKAVVLTG 660
GRGAEDAAVP AQKLRRNGIS VLVVGVGPVL SEGLRLRAGP RDSLIIHVAAY ADLRYHQDVL 720
```

IEWLCEGAQK PVNLCKPSPC MNEGSCVLQN GSYRCKCRDG WEGPHCENRE WSSCSVCVSQ 780
GWILETFLRH MAPVQEGSSR TPPSNYREGL GTEMVPTFWN VCAPGP

Seq ID NO: 442 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..2424

1 11 21 31 41 51
10 ATGCCCCCTT TCCGTGTGCT GGAGGCCGTC TGTGTTTTC TGTGTTTCCAG AGTGCCCCCA 60
TCTCTCCCTC TCCAGGAAGT CCATGTAAGC AAAGAAACCA TCGGGAAGAT TTGAGCTGCC 120
AGCAAAATGA TGTGTTGCTC GGCTGCAGTG GACATCATGT TTCTGTTAGA TGGGTCTAAC 180
AGCGTCGGGA AAGGGAGCTT TGAAGAGTCC AAGCACTTTG CCATCAGAGT CTGTGACGGT 240
CTGGACATCA GCCCCGAGAG GGTACAGATG GGAGCATTCC AGTTCAGTTC CACTCCTCAT 300
15 CTGGAATTCC CCTTGGATTG ATTTTCAACC CAACAGGAAG TGAAGGCAAG AATCAAGAGG 360
ATGGTTTTCG AAGGAGGGCG CACGGAGACG GAACCTTGCTC TGAATACCT TCTGCACAGA 420
GGGTGCGCTG GAGGCAGAAA TGCTTCTGTG CCCCAGATCC TCATCATCGT CACTGATGGG 480
AAGTCCACAG GGGATGTGGC ACTGCCATCC AAGCAGCTGA AGGAAAGGGG TGTCACTGTG 540
TTTGCTGTGG GGGTCAGGTT TCCAGGTGG GAGGAGCTGC ATGCACTGGC CAGCGAGCCT 600
20 AGAGGGCAGC ACCTGCTGTT GGCTGAGCAG GTGGAGGATG CCACCAACGG CCTCTTCAGC 660
ACCTCAGACA GCTCGGCCAT CTGCTCCAGC GCCACGCCAG ACTGCAGGGT CGAGGCTCAC 720
CCCTGTGAGC ACAGGACGCT GCAGATGGTC CGGAGATTGG CTGGCAATGG CCCATGCTGG 780
AGAGGATGCG GCGGACCCCT TGCGGTGCTG GCTGCACACT GTCCCTTCTA CAGCTGGAAG 840
AGAGTGTTC TAACCCACCC TGCCACCTGC TACAGGACCA CCTGCCCAGG CCCCTGTGAC 900
25 TCGCAGCCCT TCGCAGAAATG AGGCACATGT GTTCCAGAAG GACTGGACGG CTACCAAGTGC 960
CTCTGCCCCG TGGGCTTTGG AGGGGAGGCT AACTGTGCCC TGAAGCTGAG CTGGAATGC 1020
AGGGTGCAGC TCCTCTTCCT GCTGGACAGC TCTGCGGGCA CCACTCTGGA CGGCTTCTGT 1080
CGGGCCAAAG TCTTGTGTAA GCGGTTTGTG CCGGCGGTGC TGAGCGAGGA CTCTCGGGCC 1140
CGAGTGGGTG TGGCCACATA CAGCAGGGAG CTGCTGGTGG CGGTGCTGTG GGGGGAGTAC 1200
30 CAGGATGTGC CTGACCTGTT CTGGAGCCTC GATGGCATTC CCTTCGGTGG TGGCCCCACC 1260
CTGACGGGCA GTGCTTGGG GCAGGCGGCA GAGCGTGGCT TCGGAGCGCG CACCAAGACA 1320
GGCCAGGAGC GGCACAGTAG AGTGGTGGTT TTGCTCACTG AGTCACACTC CGAGGATGAG 1380
GTTGCGGGCC CAGCGCTCA CGCAAGGGCG CGAGAGCTGC TCCTGCTGGG TGTAGGCACT 1440
GAGGCGGTGC GGGCAGAGCT GGAGGAGATC ACAGGACGCC CAAAGCATGT GATGCTCTAC 1500
35 TCGGATCCTC AGGATCTGTT CAACCAATC CCTGAGCTGC AGGGGAAGCT GTGCAGCCGG 1560
CAGCGGCCAG GGTGCGGAGC ACAAGCCCTG GACCTGTGCT TCATGTTGGA CACCTCTGCC 1620
TCAGTAGGGC CCGAATATT TGCTCAGATG CAGAGCTTTG TGAGAAGCTG TGCCCTCCAG 1680
TTTGAGGTGA ACCCTGACGT GACACAGGTC GGCCTGGTGG TGTATGSCAG CCAGGTGCAG 1740
ACTGCTTGG GGTGAGCAC CAAACCCACC CGGGCTCGGA TGCTGGGGCC CATTAGCCAG 1800
40 GCCCCTGAGT TAGTGGGGT GGGCTCAGCC GGCACCGCCC TGCTGCACAT CTATGACAAA 1860
GTGATGACCG TCCAGAGGGG TGCCCGGCGT GGTGTCCCCA AAGCTGTGGT GGTGCTCACA 1920
GGCGGGAGG GCGCAGAGA TGCAGCGGTT CCGCCAGAGA AGCTGAGGAA CAATGGCATC 1980
TCTGCTTGG TCGTGGGCGT GGGGCTGTC CTAAGTGAGG GTCTGCGGAG GCTTGCAAGT 2040
CCCCGGGATT CCCTGATCCA CTGGGAGCT TACGCGAGCC TGCGGTACCA CCAGGACGTT 2100
45 CTATTGAGT GGTGTGTTGG AGAAGCCAA GAGCCAGTCA ACCTCTGCAA ACCAGCCCG 2160
TGATGAAAT AGGGAGCTG CTCTCTGAG AATGGAGCT ACCGCTGCAA GTGTGGGAT 2220
GGCTGGAGG GCGCCCACTG CGAGAACCGT GAGTGGAGCT CTTGCTCTGT ATGTGTGAGC 2280
CAGGATGGA TTCTTGAGAC GCCCCTGAGG CACATGGCTC CGGTGACGGA GGGCAGCAGC 2340
50 CGTACCCTTC CCAGCAACTA CAGAGAAGGC CTGGGCACTG AAATGGTGCC TACCTTCTGG 2400
AATGCTGTGG CCCAGGTGCC TTAG

Seq ID NO: 443 Protein sequence
Protein Accession #: Eos sequence

1 11 21 31 41 51
55 MPPFLLEAV CVFLPSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLDGSN 60
SVKGSFERS KHPAIVCDG LDISPERVRV GAFQFSSTPH LEFPLDSFST QQEVKARIKR 120
60 MVFKGGRTET ELALKYLLHR GLPGGRNASV PQILITVTDG KSQGDVALPS KQLKRGVTV 180
FAVGVRFPFW EELHALASEP RGQHVLLAEQ VEDATNGLPS TLSSSAICSS ATPDCRVEAH 240
PCEHRTLENV REFAGNAPCW RGSRTLAVAL AAHCFYFWSK RVFLTHPATC YRTTCFPCPD 300
SQPQNGTCT VPBGIDGYQC LCPLAFGGEA NCALRLSLEC RVDLLFLDLS SAGTTLDGFL 360
RAKVPVRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPLVWSL DGIPFRGGPT 420
LTGSALRQAA ERGPGSATRT QQDRPRRVVV LLTESHSEDS VAGPARHARA RELLLLGVGS 480
65 EAVRAELEEI TGSPIHVMVY SDPQDLFNQI PELQKLCSS QRPQCRTQAL DLVFMIDTSA 540
SVGPENFAOM QSFVRSALQ FEVNEVDVQV GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ 600
APYLGVGSA GTALLHIYDK VMTVQRGARP GVPKAVVVLT GGRGAEDAAV PAQKLNNNGI 660
SVLVGVGVFV LSEGLRRLAG PRDSLHVAA YADLRYHQDV LIENLOGEAK QPVNLCKPSP 720
70 CMNEGSCVLQ NGSYRCKCRD GWEGPHCENR ENSSCSVCVS QGWILETFLR HMAPVQEGSS 780
RTPPSNYREGL GTEMVPTFW NVCAPGP

Seq ID NO: 444 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 89..2356

1 11 21 31 41 51
75 GCCCCTGGC CCGAGCCGCG CCGGGTCTG TGAGTAGAGC CGCCCGGGCA CCGAGCGCTG 60
GTGCGCGCTC TCCTCCCGTT ATATCAACAT GCCCCTTTC CTGTTGCTGG AAGCGGTCTG 120
80 TGTITTCCTG TTTTCCAGAG TGCCCCATC TCTCCTCTC CAGGAAGTCC ATGTAAGCAA 180
AGAAACCATC GGAAGATTG CAGCTGCCAG CAAATGATG TGGTGTCTGG CTGCACTGGA 240
CATCATGTTT CTGTAGATG GGTCTAACAG CGTCGGGAAA GGGAGCTTTG AAAGGTCCAA 300
GCACCTTGGC ATCAGAGTCT GTGACGGTCT GGACATCAGC CCGAGAGGGG TCAGAGTGGG 360
85 AGCATTCCAG TTCAGTCCA CTCTCATCT GGAATCCCC TTGATTTCAT TTTCAACCCA 420
ACAGGAAGTG AAGGCAAGAA TCAAGAGGAT GGTTTTCAA GAGGGGCGCA CGGAGACGGA 480
ACTTGTCTCTG AAATACCTTC TGACAGAGG GTTGCTTGA GCAGAAATG CTTCTGTGCC 540
CCAGATCCTC ATCATGCTCA CTGATGGGAA GTCCAGGGG GATGTGGCAC TGCCATCCAA 600

	GCAGCTGAAG	GAAGGGGGTG	TCACGTGTGT	TGCTGTGGGG	GTCAGGTTTC	CCAGGTGGGA	660
	GGAGCTGCAT	GCACCTGGCCA	GGGAGCCTAG	AGGGCAGCAC	GTGCTGTTGG	CTGAGCAGGT	720
	GGAGGATGCC	ACCAACGGCC	TCTTCAGCAC	CCTCAGCAGC	TGGCCATCTT	AGTCCAGCGC	780
5	CACGCCAGAG	TGCGGGTGG	AGGCTCACCC	CTGTGAGCAC	AGGACGCTGG	AGATGGTCCG	840
	GGAGTTCGCT	GGCAATGCC	CATGCTGGAG	AGGATCGGG	CGGACCCCTG	CGGTGCTGGC	900
	TGCACACTGT	CCCTTCTACA	GCTGGAAGAG	AGTGTTCCTA	ACCCACCCCTG	CCACCTGTCTA	960
	CAGGACCACT	TGCCCAGGCG	CCTGTGACTC	GCAGCCCTGC	CAGAAATGGAG	GCACATGTGT	1020
	TCCAGAAAGGA	CTGGACGGCT	ACCAGTGCTT	CTGCCCCGCTG	GCCTTTGGAG	GGGAGGCTAA	1080
10	CTGTGCCCTG	AAGCTGAGCC	TGGAATGCAG	GGTGCACCTC	CTCTTCTGTC	TGGACAGCTC	1140
	TGCGGSCACC	ACTCTGGAAG	GCTTCTGCG	GGCCAAAGTC	TTGCTGAAGC	GGTTTGTGCG	1200
	GGCCGTGCTG	AGCGAGGACT	CTCGGGCCCG	AGTGGGTGTG	GCCACATACA	GCAGGGAGCT	1260
	GCTGTGGGCG	GTGCGCTGTG	GGGAGTACCA	GGATGTGCTT	GACCTGGTCT	GGAGCCTCGA	1320
	TGGCATTCCC	TTCCGTGGTG	GGCCACCCCT	GACGGGCAGT	GCCTTGGCGC	AGGCGGCAGA	1380
	CGTGGCTTC	GGGAGCGCCA	CCAGGACAGG	CCAGGACCGG	CCACGTAGAG	TGGTGGTTTT	1440
15	GCTCACTGAG	TCACACTCCG	AGGATGAGGT	TGCGGGCCCA	GCCTGTCACG	CAAGGGCGCG	1500
	AGAGCTGCTC	CTGCTGGGTG	TAGGCACTGA	GGCCGTGCGG	GCAGAGCTGG	AGGAGATCAC	1560
	AGGCAGCCCA	AAGCATGTGA	TGGTCTACTC	GGATCTCAG	GATCTGTTCA	ACCAATATCC	1620
	TGAGCTGCAG	GGGAAGCTGT	GCAGCCGCGA	CGCGCCAGGG	TGCGGCACAC	AAGCCCTGGA	1680
20	CCTCGTCTTC	ATGTTGGACA	CCTCTGCCTC	AGTAGGGCCC	GAGAAATTTG	CTCAGATGCA	1740
	GAGCTTTGTG	AGAAGCTGTG	CCCTCCAGTT	TGAGGTGAAC	CCTGACGTGA	CACAGGTGCG	1800
	CCTGTGGGTG	TATGGCAGCC	AGGTGCAGAC	TGCTTTCGGG	CTGGACACCA	AACCCACCCG	1860
	GGCTGCGATG	TGCGGGGCGA	TTAGCCAGGC	CCCCTACCTA	GGTGGGGTGG	GCTCAGCCGG	1920
	CACCGCCCTG	CTGCACATCT	ATGACAAAGT	GATGACCGTC	CAGAGGGGTG	CCCGGCGCTG	1980
25	TGTCCCAAAA	CGCTGTGGTG	TGCTCACAGG	CGGGAGAGGC	GCAGAGGATG	CAGCCGTTCC	2040
	TGCCCAAGAG	CTGAGGAACA	ATGGCATCTC	TGTCTTGGTC	TGGGGCGTGG	GGCCTGTCTT	2100
	AAGTGAGGGT	CTGCGGAGGC	TTGCAGGTCC	CCGGGATTC	CTGATCCACG	TGGCAGCTTA	2160
	CGCCGACCTG	CGGATACACC	AGGACGTGCT	CATTGAGTGG	CTGTGTGGAG	AAGCCAAAGCA	2220
	GCAGTCAAC	CTCTGCAAAC	CCAGCCCGTG	CATGAATGAG	GGCAGCTGCG	TCCTGCAGAA	2280
	TGGGAGCTAC	CCCTGCAAGT	GTGCGGATGG	CTGGAGGGGC	CCCCACTGCG	AGAACCAGTT	2340
30	CTTGAGACGC	CCTTGAGGCG	CATGGCTCCC	GTGAGGAGG	GCAGCAGCGG	TACCCCTCCC	2400
	AGCAACTACA	GAGAAGCGCT	GGGCACTGAA	ATGGTGCCTA	CCTTCTGGAA	TGTCTGTGCC	2460
	CCAGGTCCCT	AGAATGTCTG	CTTCCCGCCG	TGGCCAGGAC	CACATTTCTC	ACTGAGGGAG	2520
	GAGGATGCC	CACTGACAGC	CATGCTGCTT	AGAGACAAGA	AAGCAGCTGA	TGTACCCAC	2580
35	AAACGATGTT	GTTGAAAAGT	TTTGATGTGT	AAGTAAATAC	CCACTTTCTG	TACCTGCTGT	2640
	GCCTTGTGTA	GGCTATGTCA	TCTGCCACCT	TTCCCTTGAG	GATAAACAAAG	GGGTCTTGAA	2700
	GACTTAAATT	TAGCGGCGTG	ACGTTCCCTT	GCACACAATC	AATGCTCGCC	AGAATGTTGT	2760
	TGACACAGTA	ATGCCAGCA	GAGGCCTTTA	CTAGAGCATC	CTTTGAGCGG		

Seq ID NO: 445 Protein sequence
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
45	MPPFLLEAV	CVFLPSRVPP	SLPLQEVHVS	KETIGKISAA	SKMMWCSAAV	DIMFLLDGSN	60
	SVGKGSFERS	KHFAITVCDG	LDISPERVRV	GAFQPSSTPH	LEFLDPSFT	QQEVKARIKR	120
	MVFKGGRTET	ELALKYLLHR	GLPGGRNASV	PQILIIIVTDG	KSQGDVALPS	KQLKERGVTV	180
	FAVGVRFPFR	EBLHALASEP	RQHVLLAEQ	VEDATNGLPS	TLSSSAICSS	ATPDCRVEAH	240
	PCEHRTLEMV	REFAGNAPCW	RSSRRTLAVL	AAHCPFYSWK	RVFLTHPATC	YRTTCPGPCD	300
	SQPCQNGGTC	VPEGLDGYQC	LCPLAFGGEA	NCKALKSLEC	RVDLLFLDLS	SAGTTLDGFL	360
50	RAKVFVRFPV	RAVLSEDSRA	RVGVATYSRE	LLVAVPVGEY	QDVPLDVWSL	DGIPFRGGPT	420
	LTGSALRQAA	ERGFSGSATRT	GDPRPRRVVV	LLTSHSEDE	VAGPARHARA	RELLLLGVGS	480
	EAVRAELERI	TGSPKHYMVVY	SDPDLFNQI	PELQGLCSR	QRPGCRTQAL	DLVFMLDSTA	540
	SVGPENFAQM	QSFVRSALQ	FEVNPEDVTQV	GLVVYGSQVQ	TAFGLDTKPT	RAAMLRAISQ	600
	APYLGGVGA	GATLLHYDK	VMTVQRGARP	GVPKAVVLT	GGRGAEAAV	PAQKLNNNGI	660
55	SVLVVGVGPV	LSEGLRLLAG	PRDSLHVAA	YADLRHYQDV	LIEWLCEGAK	QPVNLCKPSP	720
	CNNEGSCVLQ	NGSYRCKCRD	GWEGPHCENR	FLRRP			

Seq ID NO: 446 DNA sequence
Nucleic Acid Accession #: NM_031942.1
Coding sequence: 145..1260

	1	11	21	31	41	51	
65	CCCGAGCCCC	GCCCTCCCG	GCCCGGGTGG	GCGCGCCAG	CCTGCCAGCC	GCGCTGCTGC	60
	TGCTOCTCCT	GCTGTGGGAC	CGCTGACCGC	GCGGCTGCTC	CGCTCTCCCC	GCTCCAAGCG	120
	CCGATCTGGG	CACCCGCCAC	CAGCATGGAC	GCTCGCGCGG	TGCGCGAGAA	AGATCTCAGA	180
	GTAAGAAGA	ACTTAAAGAA	ATTGAGATAT	GTGAAGTTGA	TTTCCATGGA	AACCTCGTCA	240
	TCCTCTGATG	ACAGTTGTGA	CAGCTTTGCT	TCTGATAATT	TTGCAAAACAC	GAGGCTGCAG	300
	TCAGTTGGGG	AAGGCTGTAG	GACCCGAGC	CAGTGCAAGC	ACTCTGGACC	TCTCAGGGTG	360
70	GGATGAAGT	TTCCAGCGCG	GAGTACCAGG	GGAGCAACCA	ACAAAAAAGC	AGAGTCCCGC	420
	CAGCCCTCAG	AGAATCTCTG	GACTGATTCC	AACTCCGATT	CAGAAAGATGA	AAGTGGAAATG	480
	AATTTTTTGG	AGAAAAAGGCT	TTTAAATATA	AAGCAAAACA	AAGCAATGCT	TGCAAAACTC	540
	ATGCTGTAAT	TAGAAAGCTT	CCCTGGCTCG	TTCCGTGGAA	GACATCCCCC	CCCAGGCTCC	600
	GACTCACAA	CAAGGAGACC	GCGAAGGCGT	ACATTCGCGG	GTGTTGCTTC	CAGGAGAAAC	660
75	CCTGAACGGA	GAGCTCGTCC	TCTTACCAGG	TCAAGGTCCC	GGATCCTCGG	GTCCCTTGAC	720
	GCTCTACCCA	TGGAGGAGGA	GGAGGAAGAG	GATAAGTACA	TGTTGGTGAG	AAAGAGGAAG	780
	ACCGTGGATG	GCTACATGAA	TGAAGATGAC	CTGCCACAGAA	GCCGTCGCTC	CAGATCATCC	840
	GTGACCCCTC	GACCATATAAT	TGCGCCAGTG	GAAGAAATTA	CAGAGGAGGA	GTTGGAGAAC	900
	GTCTGCAGCA	ATTCTCGAGA	GAAGATATAT	AACCGTTTAC	TGGGCTCTAC	TTGTCTATCA	960
80	TGCGCTCAGA	AGACTATTTGA	TACCAAAACA	AACCTGCAGAA	ACCCAGACTG	CTGGGGCGTT	1020
	CGAGGCCAGT	TCTGTGCGCC	CTGCCCTCGA	AACCGTTATG	GTGAAGAGGT	CAGGGATGCT	1080
	CTGCTGGATC	CGAACTGGCA	TGCGCCGCTT	TGTCGAGGAA	TCTGCAACTG	CAGTTTCTGC	1140
	CGGACGCGAG	ATGACGCGTG	TGCGACTGGG	GTCCCTTGTG	ATTTAGCCAA	ATATCATGGC	1200
	TTTGGGAATG	TGCAATGCTTA	CTTGAAAGGC	CTGAAACAGG	AATTTGAAAT	GCAAGCATAA	1260
85	TATCTGGAAT	ATTGTGCTGC	TGCCCTTCTAC	TCTCAAAATC	TTTCTGTGAA	AAGTTTCCAA	1320
	TTTTTTCACT	GAACCTGTAG	TTAAAAATCT	TGATGATCAG	CCTGTTTCAT	AAGAACTCC	1380
	AATCAAGTTA	ATCTTAGCAG	ACATGTGTTT	CTGGAGCATC	ACAGAAGGTA	TATTGCTAGT	1440

5
10
15
20

TACACCTTGC CCTCCTGCAG TTTCTTCTCT GCTCCCAACC CCCATCTCAT AGCATCCCCC 1500
TCTATTCCA ATGCTCTCTT CCAACCGCTT AGTTTCTGAA TTTCTTTTAA ATTACAGTTT 1560
TATGAAGCA TATTTTATTT ACTTGGTGTT GAAATAGCCC TCATAAAACC TAAGCAGCTG 1620
GAAACACAAT AATAGTATTA ACTAACTAGA TCTATTGAAT TTCAGAGAAG AGCCTTCTAA 1680
CTTGTTTACA CAAAACGAG TATGATTAG CACTCATACT AGTTGAAATT TTTAATAGAA 1740
TCAAGGCACA AAGGTCTTAA AACCATGTGG AAAAATTAGG TAATTATTGC AGATTGATGT 1800
CTCTCAATCC CATGTATTGC GCTTATGTTA CAAGTTGTTG TCACAGTTGA GACTTAATTT 1860
CTCTCAATTT CTTCGCCCC AAGGGTAAGT GGTGCGTCCA GCTTACACGA TCATAATTC 1920
AAGGTTGGTG GGCATGTAA TACTTAATTA AAATAATGAT GGAAGAGCTA TCTGGAGATT 1980
ATGAGTAAGC TGATTTGAAT TTTCACTATA AAACCTTAGT ATAATTGTAG TTTGCAAGT 2040
TTATTTCAGT TCACATGTAA GGTATTGCAA ATAAATTCCT GGACAATTTT GTATGGAAAC 2100
TTGATATTAA AAACATAGTT GTGGTTCTTT GCAGTTTCTT GTAAATTTAT AAACAGGCA 2160
CAAGGTTCAA GTTTAGATTT TAAGCACTTT TATAACAATG ATAAGTGCCT TTTTGAGAT 2220
GTAACCTTTA GCAGTTTGTG AACCTGACAT CTCTGCCAGT CTAGTTTCTG GGCAGGTTTC 2280
CTGTGTCAGT ATTCCCTCCT CTCTTTCAT TAATCAAGGT ATTTGGTAGA GGTGGAATCT 2340
AAGTGTGTGT ATGTCCAAAT TACTTGCTATA TGTAACCAT TGCTGTGCCA TTCAATGTTT 2400
GATGCATAAT TGGACCTTGA ATCGATAAGT GTAAATACAG CTTTGTATCT GTAATGCTTT 2460
TATACAAAG TTTATTTTAA TAATAAAATG TTTGTTCTAA AAAAAAATA

Seq ID NO: 447 Protein sequence
Protein Accession #: NP_114148.1

25
30

1 11 21 31 41 51
MDARRVPQKD LRVKQNLKPF RYVKLISMET SSSDDSCDS FASDNFANTR LQSVREGCRT 60
RSQCRHSGPL RVAMKFPARS TRGATNKAE SRQPSNSVT DSNDSSEDES GNMFLKRAL 120
NIKQNKAMLA KLMSELESFP GSFRGRHPLP GSDSQSRRPR RRTFPGVASR RNPERRARPL 180
TRSRRIILGS LDALPMEEEB EEDKYMLVRK RKTVDGYMNE DDLPRSRRSR SSVTLPHIIR 240
PVBETIEEEL ENVCNSNREK IYNRSLGSTC HQCRQKTI DT KTNCRNPDCW GVRGQFCGPG 300
LRNRYGEEVR DALDPNWHC PPCRGICNCS FCRQRDGRCA TGVLYLAKY HGFQNVHAYL 360
KSLKQEFEMQ A

Seq ID NO: 448 DNA sequence
Nucleic Acid Accession #: NM_019894
Coding sequence: 1..1314

35
40
45
50
55
60

1 11 21 31 41 51
ATGTTACAGG ATCCCTGACAG TGATCAACCT CTGAACAGCC TOGATGTCAA ACCCTGCGC 60
AAAACCCGTA TCCCATGGA GACCTTCAGA AAGGTGGGGA TCCCATCAT CATAGCACTA 120
CTGAGCTCGG CGAGTATCAT CATTTGGGTT GTCCCTCATCA AGGTGATTCT GGATAAATAC 180
TACTTCTCTT CGGGCAGACC TCTCCACTTC ATCCGAGGA AGCAGCTGTG TGACGGAGAG 240
CTGGAAGTGC CTTGGGGGA GGAAGAGGAG CACTGTGTCA AGAGCTTCCC CGAAGGGCCT 300
GCAGTGGCAG TCCGCTCTC CAAGGACCGA TCCACACTGC AGGTGCTGGA CTCGGCCACA 360
GGGAACCTGG TCTCTGCTG TTTGACAAAC TTCACAGAAG CTCTCGCTGA GACAGCCTGT 420
AGGCAGATGG GCTACAGCAG CAAACCCACT TTCAGAGCTG TGGAGATTGG CCCAGACCAG 480
GATCTGGATG TTGTTGAAT CACAGAAAC AGCCAGGAGC TTGCGATGCG GAACTCAAGT 540
GGGCCCTGTC TCTCAGGCTC CTGCTCTCC CTGCACTGTC TTGCTGTGG GAAGAGCCTG 600
AAGACCCCCC GTGTGGTGGG TGGGGAGGAG GCCTCTGTGG ATTCTTGGCC TTGGCAGGTC 660
AGCATCCAGT AGACAAACA GCACGTCTGT GGAGGGAGCA TCCTGGACCC CCACTGGGTC 720
CTCAGGCGAG CCCACTGCTT CAGGAACCAT ACCGATGTGT TCAACTGGAA GGTGCGGGCA 780
GGCTCAGACA AACTGGGCG CTTCCCATCC CTGGCTGTGG CCAAGATCAT CATCATTGAA 840
TTCAACCCCA TGTACCCCAA AGACAATGAC ATCGCCCTCA TGAAGCTGCA GTTCCCCTC 900
ACTTTCTCAG GCACAGTCAG GCCCATCTGT CTGCCCTTCT TTGATGAGGA GCTCACTCCA 960
GCCACCCAC TCTGGATCAT TGGATGGGGC TTTACGAAGC AGAATGGAGG GAAGATGTCT 1020
GACATCTGCG TCCAGGCTGC AGTCCAGGTC ATTGACAGCA CACGGTGCAA TGCAGACGAT 1080
GCGTACCAGG GGGAAAGTCA CGAGAAGATG ATGTGTGCG GCATCCCGGA AGGGGGTGTG 1140
GACACCTGCC AGGGTGACAG TGGTGGGCCC CTGATGTACC AATCTGACCA GTGGCATGTG 1200
GTGGGCAATG TTAGCTGGGG CTATGGCTGC GGGGGCCCGA GCACCCAGG AGTATACACC 1260
AAGGTCTCAG CCTATCTCAA CTGGATCTAC AATGTCTGGA AGGCTGAGCT GTAA

Seq ID NO: 449 Protein sequence
Protein Accession #: NP_063947.1

65
70
75

1 11 21 31 41 51
MLQDPDSQD LNSLDVKPLR KPRIPMETFR KVGIPILIAL LSLASIIIV VLIKVILDKY 60
YFLCGQLHP IPRKQLDGE LDCPLGEDEB HCVKSPFEGP AVAVRLSKDR STLQVLDSAT 120
GNWFSACFDN FTEALAEAC RQMGYSKPT FRAVEIGPDQ DLDVVEITEN SQELMRNNS 180
GPCLSGSLVS LHCLACGKSL KTRPVVGGEE ASVDSNPNQV SIQYDKQHV GGSILDPHNV 240
LTAABCFRKH TDVFNWVRA GSDKLGSFPS LAVAKIIIE FNPMPKDN DIALMLQFPL 300
TFSGTVRPI LPPFDEELTP ATPLNIIGWG FTKQNGGKMS DILLQASVQV IDSTRCNADD 360
AYQGEVTEK MCAGIPEGV DTCQGDGSGP LMYQSDQNHV VGIVSNWGYC GGPSTPGVYT 420

Seq ID NO: 450 DNA sequence
Nucleic Acid Accession #: XM_051860.2
Coding sequence: 52..3042

80
85

1 11 21 31 41 51
GCTCACCCAG GAAAAATATG CAATGCTCCC ATTGATATAC AGGCCACTAC AATGGATGGA 60
GTTAACCTCA GCACCGAGGT TGTCTACAAA AAAGGCCAGG ATTATAGGTT TGCTTGCTAC 120
GACCGGGGCA GAGCCTGCGG GAGCTACCGT GTACGGTTCC TCTGTGGGAA GCCTGTGAGG 180
CCCAACCTCA CAGTACCAT TGACACCAAT GTGAACAGCA CCATTCTGAA CTTGGAGGAT 240
AATGTACAGT CATGGAACC TGGAGATACC CTGGTCATTG CCAGTACTGA TTAATCCATG 300
TACCAGGCGA AAGAGTTCCA GGTGCTTCCC TGCAGATCCT GCGCCCCCAA CCAGGTCAAA 360

	GTGGCAGGGA	AACCAATGTA	CTGTGACATC	GGGGAGGAGA	TAGACGGCGT	GGACATGCGG	420
	GCGGAGGTG	GGCTTCTGAG	CCGGAACATC	ATAGTGATGG	GGGAGATGGA	GGACAAATGC	480
	TACCCCTACA	GAACCAACAT	CTGCAATTTT	TTTGACTTGG	ATACCTTTGG	GGGCCACATC	540
5	AAGTTTGCTC	TGGGATTAA	GGCAGCACAC	TTGGAGGGCA	CGGAGCTGAA	GCATATGGGA	600
	CAGCAGCTGG	TGGGTACGTA	CCGATTTCAC	TTCCACCTGG	CCGGTGATGT	AGACGAAAGG	660
	GGAGGTTATG	ACCCACCCAC	ATACATCAGG	GACCTCTCCA	TCCATCATAC	ATTCTCTCGC	720
	TGCGTCACAG	TCCAATGGCT	CAATGGCTTG	TTGATCAAGG	ACGTTGTGGG	CTATAACTCT	780
	TTGGGCCACT	GCTTCTTCAC	GGAGATGGG	CCGGAGGAAAC	GCAACACTTT	TGACCACTGT	840
10	CTTGGCCTCC	TTGTCAAGTG	TGGAACCTTC	CTCCCTCGG	ACCGTGACAG	CAAGATGTGC	900
	AAGATGATCA	CAGGAGACTC	CTACCCAGGG	TACATCCCCA	AGCCACGGCA	AGACTGCAAT	960
	GCTGTGTCCA	CCTTCTGGAT	GGCCAATCCC	AACAACAACC	TCATCAACTG	TGCCGCTGCA	1020
	GGATCTGAGG	AAACTGGATT	TTGGTTTATT	TTTCCACAG	TACCAACGGG	CCCTCCCGTG	1080
	GGAAATGTACT	CCCCAGGTTA	TTGAGAGCAC	ATTCCACTGG	GAAAAATCTA	TAACAACCGA	1140
	GCACATTCGA	ACTACCGGGC	TGGCATGATC	ATAGACAACG	GAGTCAAAAC	CACCGAGGCC	1200
15	TCTGCCAAGG	ACAAGCGGCC	GTTCTCTCA	ATCATCTCTG	CCAGATACAG	CCCTCACCAG	1260
	GAGCGCGACC	CGCTGAAGCC	CCGGAGGCGG	GCCATCATCA	GACACTTCAT	TGCCCTACAG	1320
	AACCAGGACC	ACGGGGCTCT	GCTGCGCGGC	GGGGATGTGT	GGCTGGACAG	CTGCCGTTT	1380
	GCTGACAAAT	GCATTGGCCT	GACCTTGGCC	AGTGGTGAA	CCTTCCCGTA	TGACGACGGC	1440
20	TCCAAGCAAG	AGATTAAGAA	CAGCTTGTGT	GTGGCGAGA	GTGGCAACGT	GGGGACGGAA	1500
	ATGATGGACA	ATAGGATCTG	GGGCCCTGGC	GGCTTGGACC	ATAGCGGAAG	GACCTCCCT	1560
	ATAGGGCAGA	ATTTTCCAAT	TAGAGGAATT	CAGTTATATG	ATGGCCCCAT	CAACATCCAA	1620
	AACGTGCATT	TCCGAAGTGT	TGTGGCCCTG	GAGGGCCGGC	ACACCAGCGC	CCTGGCCCTT	1680
	CGCCTGAATA	ATGCTGGGCA	GAGCTGCCCC	CATAACAACG	TGACCGGCTT	TGCTTTGAG	1740
25	GACGTTCCGA	TACTTCCAG	AGTGTCTTC	GGAGAGCCTG	GGCCCTGGTT	CAACCAGCTG	1800
	GACATGGATG	GGGATAAGAC	ATCTGTGTTT	CATGACGTCT	ACGGCTCCGT	GTCCGAGTAC	1860
	CCTGGCTCCT	ACCTCAGCAA	GAATGACAAC	TGGCTGGTCC	GGCACCCAGA	CTGCATCAAT	1920
	GTTCCCGACT	GGAGAGGGGC	CATTGTCAGT	GGGTGCTATG	CACAGATGTA	CATTCAAGCC	1980
	TACAAGACCA	GTAACCTGGG	AATGAAGATC	ATCAAGAATG	ACTTCCCGAG	CCACCTCTCT	2040
30	TACCTGGAGG	GGCGCTCAC	CAGGAGCACC	CATTACCAGC	AATACCAACC	GGTTGTCAAC	2100
	CTGCAGAAAG	GCTCAACCAT	CCACTGGGAC	CAGACGGCCC	CCGCGGAATC	CGCCATCTGG	2160
	CTCATCAACT	TCAACAAGGG	CGACTGGATC	CGATGGGGGC	TCTGTACCC	GGGAGGCACC	2220
	ACATTTCTCA	TCTCTCGGA	TGTTCACAAT	CGCCTGCTGA	AGCAAACTGC	CAAGACGGGC	2280
	GTCTTCTGTA	GGACCTTGCA	GATGGACAAA	GTGGAGCAGA	GCTACCTCTG	CAGGAGCCAC	2340
35	TACTACTGGG	ACGAGGACTC	AGGGCTGTTG	TTCTGGAAGC	TGAAAGCTCA	GAACGAGAGA	2400
	GAGAAGTTTG	CTTTCTGCTC	CATGAAAGGC	TGTGAGAGGA	TAAAGATTAA	AGCTCTGATT	2460
	CCAAAGAAAG	CAGGCGTCAG	TGACTGCACA	GCCACAGCTT	ACCCCAAGTT	CACCGAGAGG	2520
	GCTGTGCTAG	ACGTGCCGAT	GCCCAAGAA	CTCTTTGGTT	CTCAGCTGAA	AACAAGGAC	2580
	CATTCTCTGG	AGGTGAAGAT	GGAGAGTTCC	AAGCAGCACT	TCTTCCACCT	CTGGAACGAC	2640
40	TTCCGTTACA	TTGAAGTGGA	TGGGAAGAAG	TACCCCAAGT	CGGAGGATGG	CATCCAGGTG	2700
	GTGGTGATTG	ACGGGAACCA	AGGGCGCGTG	GTGAGCCACA	CGAGCTTCAG	GAACCTCAAT	2760
	CTGCAAGGCA	TACCATGGCA	GCTTTTCAAC	TATGTGGCGA	CCATCCCTGA	CAATTCCATA	2820
	GTGCTTATGG	CATCAAGGG	AAGATACGTC	TCCAGAGGCC	CATGGACCAG	AGTGTCTGAA	2880
	AAGCTTGGGG	CAGACAGGGG	TCTCAAGTTG	AAAGAGCAAA	TGGCATTCTG	TGGCTTCAAA	2940
45	GGCAGCTTCC	GGCCCATCTG	GGTGACACTG	GACACTGAGG	ATCACAAGGC	CAAAATCTTC	3000
	CAAGTTGTGC	CCATCCCTGT	GGTGAAGAA	AAGAAGTTGT	GAGGACAGCT	GGCGCCCGGT	3060
	GCCACCTCGT	GGTAGACTAT	GACGGTGACT	CTTGGCAGCA	GACCAAGTGG	GGATGGCTGG	3120
	GTCCCCCAGC	CCCTGCCAGC	AGCTGCCCTG	GAAGGCGGTG	TTTCAGCCCT	GATGGGCCAA	3180
	GGGAAGGCTA	TCAGAGACCC	TGGTGCTGCC	ACCTGCCCTT	ACTCAAGTGT	CTACCTGGAG	3240
50	CCCTCGGGGC	GGTGCTGGCC	AATGCTGGAA	ACATTCACTT	TCTTGCAGCC	TCTTGGGTGC	3300
	TTCTCTCTTA	TTCTTGCTCT	TTCAAGTGGG	GTTTGGGGAC	CATATCAGGA	GACCTGGGTT	3360
	GTGCTGACAG	CAAGATTCGA	CTTTGGCAGG	AGCCCTGACC	CAGCTAGGAG	GTAGTCTGGA	3420
	GGGCTGGTGA	TTCAAGATCT	CCCATGGTCT	TCAGCAGACA	AGTGAGGGTG	GTAATGTAG	3480
	GAGAAAGAGC	CTTGCCCTTA	AGGAAATCTT	TACTCTCTGA	AGCAAGAGCC	AACCTCACAG	3540
55	GATTAGGAGC	TGGGGTAGAA	CTGGCTATCC	TTGGGGAAGA	GGCAAGCCCT	GCCTCTGGCC	3600
	GTGTCCACCT	TTGAGGAGAC	AGGTTTGGAC	TTGAGTGA	TGAGCTCTCA	TGACTCTCAA	3660
	AGGCCCTTTT	AGTTCTGAGA	TTCCAGAAAT	CTGCTGCATT	TCACATGGTA	CCTGGAACCC	3720
	AACAGTTTAT	GGATATCCAC	TGATATCCAT	GATGCTGGGT	GCCCGAGCGC	ACAAGGGATG	3780
	GAGAGGTGAG	AACTAATGCC	TAGCTTGAGG	GGTCTGCGAT	CCAGTAGGGC	AGGCAGTCAG	3840
60	GTCCATGTGC	ACTGCAATGC	CAGGTGGAGA	AATCACAGAG	AGGTAAAATG	GAGGCCAGTG	3900
	CCATTTTACA	GGGGAGGCTT	CTTGCTTACA	GGAAATGAAG	CTGGGGGCAT	CTGGGGGCAT	3960
	TTTGCTGGGG	GGAGATGAGG	CAGCCTCTGG	AATGGCTCAG	GGATTACGCC	CTCCCTGCGG	4020
	CTGCTGCTG	AGCTTGGTGA	CTACGGGGTC	GCCCTTTGCT	CACGCTCTCT	TGGCCCACTC	4080
	ATGATGGAGA	AGTGTGGTCA	GAGGGGAGCA	ATGGGCTTTG	CTGCTTATGA	GCACAGAGGA	4140
65	ATTCACTCCC	CAGGACGCCC	TGCCCTTGAC	TCCAAGAGGG	TGAAGTCCAC	AGAAGTGAGC	4200
	TCTTGCTCTA	GGGCTCTATT	TGCTCTTCAT	CCAGGGAATC	GAGCACAGGG	GGCCTCCAGG	4260
	AGACCCCTAGA	TGTGCTCGTA	CTCCCTCGGC	CTGGGATTTT	AGAGCTGGAA	ATATAGAAAA	4320
	TATCTAGCCC	AAAGCCTTGA	TTTAAACAGA	TGGGAAAGT	GAGCCCCCAA	GATGGGAAAG	4380
	AACCACACAG	CTAAGGGAGG	GCTTGGGGAG	CCCCACCTTA	GCCCTTGTCT	CCACACACCA	4440
70	TTGCCTCAAC	AACCGGCCCC	AGAGTGCCCC	GGCACTCCTG	AGGTAGCTTC	TGGAAATGGG	4500
	GACAAAGTCCC	CTCGAAGGAA	AGGAAATGAC	TAGAGTAGAA	TGACAGCTAG	CAGATCTCTT	4560
	CCCTCTCTGT	CCGAGCGCAC	ACAAACCCGC	CCTCCCTCTG	GTGTTGGCGG	TCCCTGTGGC	4620
	CTTCACTTTG	TTCACTACTG	GTGAGCCGAG	CCTGGGTGCA	CAGTAGCTGC	AACTCCCAT	4680
	TGGTGCTACC	TGGCTCTACCT	GTCTCTGACG	CTCTACAGGT	GAGGCCGAGC	AGAGGGAGTA	4740
	GGGCTCGCCA	TGTTTCTGCT	GAGCCAATTT	GGCTGATCTT	GGGTGTCTGA	ACAGCTATTG	4800
75	GGTCCACCCC	AGTCCCTTTT	AGCTGCTGCT	TAATGCCCTG	CTCTCTCCCT	GGCCCACTTT	4860
	ATAGAGAGCC	CAAGAGCTCT	CTGTAAGAGG	GAGAACTCTA	TCTGTGTTTT	ATAATCTTGC	4920
	ACGAGGCACC	AGAGTCTCCC	TGGGTCTTGT	GATGAACATC	ATTTATCCCC	TTTCTTGCCC	4980
	CAACCACAAA	CTCTTTCTCT	CAAGAGGGGC	CTGCTGGCT	CCCTCCACCC	AACTGCACCC	5040
80	ATGAGACTCG	CTCGAAGAGT	CCATTCCCCA	GGTGGGAGCC	AACTGTCAAG	GAGGTCTTTC	5100
	CCACCAAAAC	TCCTTCAGCT	GCTGGGAGGT	GACCATAGGG	CTCTGCTTTT	AAAGATATGG	5160
	CTGCTTCAAA	GGCCAGAGTC	ACAGGAAGGA	CTTCTTCCAG	GGAGATTAGT	GGTGTATGAG	5220
	AGGAGAGTTA	AAATGACCTC	ATGTCCTTCT	TGTCACAGGT	TTTGTGTAGT	TTTCACTCTT	5280
	CTAATGCAGG	GGTCTCACAC	TGTGAACACC	TTAGGATGTG	ATCACTTTCA	GGTGGCCAGG	5340
	AATGTTGAAT	GTCTTTGGCT	CAGTTCAATT	AAAAAAGATA	TCTATTGAA	AGTCTCTAGA	5400
85	GTGTATACATA	TGTTTACAG	TACAGGATCT	GTACATAAAA	GTTTCTTTCC	TAAACCATTC	5460
	ACCAAGAGCC	AATATCTAGG	CATTCTCTTG	GTAGACAAA	TTTTCTTATT	GCTTAGAAAA	5520
	TGTCTCTCT	TGTTATTCTT	GTTTGAAGA	CTTAAGTGAG	TTAGGTTCTT	AAGGAAAGCA	5580

ACGCTCCTCT GAAATGCTTG TCTTTTTTCT GTTCCCGAAA TAGCTGGTCC TTTTTCGGGA 5640
 GTTAGATGTA TAGAGTGTGT GTATGTAAAC ATTTCTTGTA GGCATCACC A TGAACAAAGA 5700
 TATATTTTCT ATTTATTTAT TATATGTGCA CTTCAGAAG TCACTGTCCAG AGAAATAAAG 5760
 AATTGTCTTA AATGTCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 451 Protein sequence
 Protein Accession #: XP_051860.2

1 11 21 31 41 51
 MDGVNLSTEV VYKKGQDYRF ACYDRGRACR SYRVRFLCGK FVRPKLTVTI DTNVNSTILN 60
 LEDNVQSWKP GDTLVIASTF YSMYQAEFPQ VLPSCRSCAPN QVKVAGKPMY LHIGEEIDGV 120
 DMRAEVLGSL RNIIIVMGEM DKCYFYRNHI CNFFDFDTFG GHKFPALGPK AAHLBGTCLK 180
 HMQQQLVGYQ PIHFHLAGDV DERGGYDPPT YIRDLISHT FSRCTVHGS NGLLIKDVVG 240
 YNSLGHCFPT EDGPPEERNTF DHCLGLLVKS GTLLPSDRDS KMKCMITGDS YPGYIPKPRQ 300
 DCNAVSTFWM ANPNNNLINC AAAGSEETGP WFIHVVPTG PSVGMYSFGY SEHIPLGKPY 360
 MNRAHSNYRA GMIIIDNGVKT TEASAKDKRP FLSTIISARYS PHQDADPLKP REPAIRIRFI 420
 AYKNQDHLGW LRGGDLVWLS CRPADNGIGL TLASGGTFPY DDGSKQELKN SLFVGESGNV 480
 GTEMMONRIW GPGGLDHSGR TLPIGNFPI RGIQLYDGP NIQNTCFRKP VALEGRHTSA 540
 LAFRIANAWQ SCPHNNVTGI AFEDVPITSR VFPGEPPWF NQLMDGDKT SVFHDVDGSV 600
 SEYPSYLYTK NDNMLVHRPD CINVPDNRGA ICSGCYAQMY IQAYKTSNLR MKIINKDFPS 660
 HPLYLGLALT RSTHYQQYQF VVTILQKGYTI HWDQTAPAEI AIWILINFKG DWIRVGLCYP 720
 RGTTFSILSD VHNRLKQTS KTGVPVRLQ MDKVEQSYPG RSHYYWDEDS GLLPLKLAQ 780
 NEREKFAFCS MKGCERIKIK ALIPKNAGVS DCTATAYPKF TERAVVDVPM PKKLFGSQLK 840
 TKDHFLEVMK ESSKQHFFHL WNDPAYIEVD GKYPSSSDG IQVVVIDGNQ GRVVSHTSFR 900
 NSILQGIPIW LFNYVATIPD NSIVLMASKG RYVSRGPWTR VLEKLGDARG LKLKQMAPV 960
 GFKGSFRPIW VTLDTEHKA KIPQVVPPIV VKKKKL

Seq ID NO: 452 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 261..2861

1 11 21 31 41 51
 GAGCTAGCGC TCAAGCAGAG CCCAGCGCGG TGCTATCGGA CAGAGCCTGG CGAGCGCAAG 60
 CGGCGCGGGG AGCCAGCGGG GCTGAGCGCG GCCAGGGTCT GAACCCAGAT TTCCAGACT 120
 AGCTACCACT CCGCTTGCCC ACGCCCGGGG AGCTCGCGGC GCCTGGCGGT CAGCGACCAAG 180
 ACGTCCGGGG CGCTGCGCTT CCGTGGCGGC GAGGCGTGAC ACTGTCTCGG CTACAGACCC 240
 AGAGGGAGCA CACTGCCAGG ATGGGAGCTG CTGGGAGGCA GGACTTCCTC TTCAAGGCCA 300
 TGCTAGCACT CAGCTGGCTC ACTCTGACCT GCTTCCCTGG GGCCACATCC ACAGTGGCTG 360
 CTGGGTGCCC TGACCAAGAG CCTGAGTTGC AACCTGGAA CCCTGGCCAT GACCAAGACC 420
 ACCATGTGCA TATCGGCCAG GGCAGACAC TGCTGTCTAC CTCTCTGCC ACGGTCTATT 480
 CCATCCACAT CTCAGAGGGA GGCAGCTGG TCATTAAGA CCACGACGAG CCGATTGTIT 540
 TGCGAACCAG GCACATCCTG ATTGACAAAG GAGGAGAGCT GCATGCTGGG AGTGCCCTCT 600
 GCCCTTTTCA GGGCAATTTT ACCATCATTT TGTATGGAAG GGCTGATGAA GGTATTTCAGC 660
 CGGATCCTTA CTATGGTCTG AAGTACATTG GGGTTGGTAA AGGAGGCGCT CTGAGTTGC 720
 ATGGACAGAA AAGGCTCTCC TGGACATTTC TGAACAAGAC CCTTCAACCA GGTGGCATGG 780
 CAGAAAGAGG CTATTTTTTT GAAAGGAGCT GGGGCCACCG TGGAGTTATT GTTCATGTCA 840
 TCGAACCCTA ATCAGGCACA GTCATCCATT CTGACCGGTT TGACACCTAT AGATCCAAGA 900
 AAGAGAGTGA ACGTCTGGTC CAGTATTGTA ACGCGGTGCC CGATGGCAGG ATCTCTTCTG 960
 TTGCACTGAA TGATGAAGGT TCTCGAAATC TGGATGACAT GGCCAGGAAG GCGATGACCA 1020
 AATTGGGAAG CAACACCTTC CTGCACTCTG GATTTAGACA CCCTTGGAGT TTTCTAACTG 1080
 TGAAGGAAA TCCATCATCT TCAGTGAAG ACCATATTGA ATATCATGGA CATCGAGGCT 1140
 CTGCTCTGCG CCGGTATTTC AAATTGTTC AGACAGAGCA TGGCGAATAT TTCAATGTTT 1200
 CTTTGTCCAG TGAGTGGGTT CAAGACGTGG AGTGGACGGA GTGGTTCCAT CATGATAAAG 1260
 TATCTCAGAC TAAAGGTGGG GAGAAAATTT CAGACCTCTG GAAAGCTCAC CCAGGAAAAA 1320
 TATGCAATCG TCCCATTTGAT ATACAGGCCA CTACAATGGA TGGAGTTAAC CTCAGCACCG 1380
 AGGTTGTCTA CAAAAAGGC CAGGATTATA GGTTCGCTTG CTACGACCGG GGCAGAGCCT 1440
 GCCGAGCTA CCGTGTACGG TTCTCTGTG GGAAGCCTGT GAGGCCCAAA CTCACAGTCA 1500
 CCATTGACAC CAATGTGAAC AGCACCATTC TGAACCTGGA GGATAATGTA CAGTCATGGA 1560
 AACCTGGAGA TACCCTGGTC ATTGCCAGTA CTGATTACTC CATGTACCAG GCAGAAAGAT 1620
 TCCAGGTGCT TCCCTGCGA TCCTGCGCCC CCAACCGAGT CAAAGTGGCA GGGAAACCAA 1680
 TGTACCTGCA CATCGGGGAG GAGATAGACG GCGTGGACAT GCGGCGGAG GTTGGGCTTC 1740
 TGAGCCGGAA CATCATAGTG ATGGGGGAGA TGGAGGACAA ATGCTACCCC TACAGAAACC 1800
 ACATCTGCAA TTTCTTTGAC TTGATACCT TTGGGGGCGA CATCAAGTTT GCTCTGGGAT 1860
 TTAAGGCAGC ACCTTGGAG GGCACGGAGC TGAAGCATAT GGGACAGCAG CTGGTGGGTC 1920
 AGTACCGAT TCACTTCCAC CTGGCCGGTG ATGTAGACGA AAGGGGAGGT TATGACCCAC 1980
 CCACATACAT CAGGGACCTC TCCATCCATC ATACATTCTC TCGTGGCTC ACAGTCCATG 2040
 GCTCCAATGG CTTGTTGATC AAGGACGTTG TGGGCTATAA CTCCTTGGGC CACTGCTTCT 2100
 TCACGGAAGA TGGGCGGAGT GAAACGCAACA CTTTGAACCA CTGTCTTGGC CTCCTTGTC 2160
 AGTCTGGAAC CCTCTCCCC TCGGACCGTG ACAGCAAGAT GTGCAAGATG ATCAGAGAGG 2220
 ACTCTACCC AGGGTACATC CCCAAGCCCA GGCAAGACTG CAATGCTGTG TCCACCTTCT 2280
 GATGCGCCAA TCCCAACAAC AACCTCATCA ACTGTGCGGC TGCAAGATCT GAGGAACTG 2340
 GATTTTGGTT TATTTTTCAC CACGTACCAA CGGGCCCCCT CGTGGGAATG TACTCCCCAG 2400
 GTTATTGAGA GCACATTCCA CTGGGAAAAT TCTATAACAA CCGAGCACAT TCCAACTACC 2460
 GGGCTGGCAT GATCAGATCA AACCGAGTCA AAACCAACGA GGCTCTGCCC AAGGACAAAG 2520
 GSCCGTTCTT CTCAATCATC TCTGCCAGAT ACAGCCCTCA CCAGGACGCC GACCCGCTGA 2580
 AGCCCCGGGA GCGCGGCATC ATCAGACACT TCATTGCCTA CAAGAACCCG GACCAAGGGG 2640
 CTTGGCTGCG CGCGGGGAT GTGTGGCTGG ACAGCTGCCA TTTTCAGAGG GAGGCTCAGG 2700
 AAGGCTTCTT GCTTACAGGA ATGAAGGCTG GGGGCATTTT GCTGGGGGGA GATGAGGCG 2760
 CCTCTGGAAT GGCTCAGGGA TTCAGCCCTC CTGCGCGCTG CCGTCTGAAG CTGGTGACTA 2820
 CGGGGTCGCC CTTTGTCTAC GTCTCTCTGG CCCACTCATG ATGGAGAAGT GTGGTCAGAG 2880
 GGGGCAATG GGTCTTGGCT GTTATGAGCA CAGAGGAATT CAGTCCCCAG GCAGCCCTGC 2940
 CTCTGACTCC AAGAGGGTGA AGTCCACAGA AGTGAGCTCC TGCCCTAGGG CCTCATTTGC 3000
 TCTTCATCCA GGAACCTGAG CACAGGGGGC CTCCAGGAGA CCCTAGATGT GCTCGTACTC 3060
 CCTCGGCGCT GGATTTCAGA GCTGGAAATA TAGAAAATAT CTAGCCCAAA GCCTTCATTT 3120

5 TAACAGATGG GGAAGTGTAG CCCCAAGAT GGGAAAGAAC CACACAGCTA AGGGAGGGCC 3180
 TGGGGAGCCC CACCCTAGCC CTGTCTGCCA CACCACATTG CCTCAACAAC CGGCCCCAGA 3240
 GTGCCAGGCC ACTCCTGAGG TAGCTTCTGG AAATGGGGAC AAGTCCCTTC GAAGGAAAGG 3300
 10 AATGACTAG AGTAGAATGA CAGCTAGCAG ATCTCTTCCC TCCTGCTCCC AGCGCACACA 3360
 AACCCGCCCT CCCCTGGGTG TTGGCGGTCC CTGTGGCCTT CACTTTGTTC ACTACCTGTC 3420
 AGCCAGCCTT GGGTGCACAG TAGCTGCAAC TCCCCATTGG TGCTACCTGG CTCTCCTGTC 3480
 TCTGCAGCTC TACAGGTAGG GCCCAGCAGA GGGAGTAGGG CTCGCCATGT TTCTGGTGGG 3540
 CCAATTTGGC TGATCTTGGG TGTCTGAACA GCTATTGGGT CCACCCCACT CCCTTTTCAGC 3600
 15 TGCTGCTTAA TGCCCTGCTC TCTCCCTGGC CCACCTTATA GAGAGCCCAA AGAGCTCCTG 3660
 TAAGAGGGAG AACTCTATCT GTGGTTTATA ATCTTGCAAG AGGCACCAGA GTCTCCCTGG 3720
 GTCTTGTGAT GAATACATT TATCCCTTTT CCTGCCCAA CCACAAACTC TTCTCTTCAA 3780
 AGAGGGCCTG CCTGCTCTCC TCCACCCAAC TGCACCCATG AGACTCGGTC CAAGAGTCCA 3840
 TTCCCCAGGT GGGAGGCCAC TGTCAGGGAG GTCTTTCCCA CCAACATCTT TTCAGCTGCT 3900
 GGGAGGTGAC CATAGGGCTC TGCTTTTAAA GATATGGCTG CTTCAAAGGC CAGAGTCACA 3960
 20 GGAAGGACTT CTCCAGGGA GATTAGTGGT GATGGAGAGG AGAGTTAAAA TGACCTCATG 4020
 TCCTCTTGT CCACGCTTTT GTTGAGTTT CACTCTTCTA ATGCAAGGGT CTCACACTGT 4080
 GAACCACTTA GGATGTGATC ACTTTTCAGT GGGCAGGAAT GTTGAATGTC TTTGGCTCAG 4140
 TTCAATTAAA AAGATATCT ATTTGAAAGT TCTCAGAGTT GTACATATGT TTCACAGTAC 4200
 AGGATCTGTA CATAAAGTT TCTTCTCTAA ACCATTCAAC AAGAGCCAAAT ATCTAGGCAAT 4260
 25 TTTCTTGGTA GCACAAATTT TCTTATGTCT TAGAAAAATG TCCTCCTTGT TATTCTGTGT 4320
 TGTAAAGCTT AAGTAGGTTA GGTCTTTAAG GAAAGCAACG CTCCTCTGAA ATGCTTGTCT 4380
 TTTTCTGTT GCCGAAATAG CTGGTCCCTT TTCGGGAGTT AGATGTATAG AGTGTTTGTA 4440
 TGTAAACATT TCTGTAGGAC ATCACCATGA ACAAAGATAT ATTTTCTATT TATTATTAT 4500
 ATGTGACATT CAGAAGTCA CTGTGAGAGA AATAAAGAAT TGTCTTAAAT GTCATGATTG 4560
 30 GAGATGTCTT TTGCATTGCT TGAAGGGGT GTACCTAGAG CCAAGGAAAT TGGCTCTGGT 4620
 TTGAAAAAT TTTGCTGTTA TTATAGTAAA CATACAAAGG ATGTCAAAAA AAAAAA 4680
 AAAAAA AAAAAA AA

Seq ID NO: 453 Protein sequence
 Protein Accession #: Eos sequence

35 1 11 21 31 41 51
 MGAAGRQDFL FKAMLTISWL TLTCFPGATS TVAAGCPDQS PELQFVNPQH DQDHHVHIQ 60
 GKTLILLTSSA TVVSIHISEG GKLVIKDHDE PIVLRTRHIL IDNGGELHAG SALCPFPQNF 120
 TIILYGRADE GIQPPPYLGL KYIGVGKGA LELHGQKKLS WFLNKLTHP GGMABGGYFF 180
 ERSWGERGVI VHVLDPKSGT VIHSDRFDTY RSKKESERLV QYLANAVPDR ILSVAVNDEG 240
 SRNLDDMARK AMTKLGSKEF LHLGFRHPWS FLT VKGNPSS SVEDHIEYHG HRGSAAARVF 300
 40 KLPTHEHGY FNVLSSEWV QDVWTEWFD HDKVSQTKGG EKISDLWKAH PGKICNRPID 360
 IQATTMDGVN LSTEVVYKKG QDYRFACYDR GRACRSYRVR FLCGKPVPRK LVTITDITNVN 420
 STILNLEDNV QSWKPGDTLV IASTDYSMYQ AEEFQVLPGR SCAPNQVKVA GKPMYLHIGE 480
 EIDGVDMRAE VGLLSRNIV MGEMEDKCP YRNHICNFFD PDTFGGHKIF ALGPKAAHLE 540
 GTELKHEMQQ LVQGYPIHFP LAGDVDERGG YDPPYIRDL SIHHTFSRCV TVHGSNGLLI 600
 45 KDVVVNSIG HCFPTEDGPE ERNTFDHCLG LLVKSGLTLP SDRDSKMKCM ITEDSYPGYI 660
 PKPRQDCNAV STFWMANPNN NLINCAAAGS EETGFWFIFH HVPTGPSVGM YSPGYSEHIP 720
 LGKPYNNRAH SNYRAGMIID NGVKTEASA XDKRPFLLSI SARYSPHQDA DPLKPREPAI 780
 IRHFYAKNQ DHGAWLRGGD WVLDSCHFRG EAQBGFLLTG MKAGGILLGG DEAAAGMAQG 840
 FSPPCRLLK LVTTGSPFAH VSLAHS

Seq ID NO: 454 DNA sequence
 Nucleic Acid Accession #: NM_013282.2
 Coding sequence: 85..2466

55 1 11 21 31 41 51
 CGACTCCTTA GAGCATGGCA TGGCTCAGAG GTGCTGGTAA AACTGATGGG GGTTTTGTCT 60
 TCCCTCCCTC TCAGGCGCGA CACCATGTGG ATCCAGGTTT GGACATGGA CGGAGGCGAG 120
 ACCCACAAGG TGAAGCTGGT GTCCAGGCTG ACCAAGGTGG AGGAGCTGAG CGGGAAGATC 180
 60 CAGGAGCTGT TCCAGTGGTA GCCAGGCTGT CAGAGGCTGT TCTACAGGGG CAAACAGATG 240
 GAGGAGCGGC ATACCTCTTT CGACTACGAG GTCCGCTGTA ATGACACCAT CCAGCTCCTG 300
 GTCCGCGAGA GCCTCGTGCT CCCCACAGC ACCAAGGAGC GGGACTCCGA GCTCTCCGAC 360
 ACCGACTCCG GCTGCTGGCT GGGCCAGAGT GAGTCAGACA AGTCTCCAC CCACGCGGAG 420
 GCGGCGCGCG AGACTGACAG CAGGCCAGCC GATGAGGACA TGTGGGATGA GACGGAATTG 480
 65 GGGCTGTACA AGGTCAATGA GTACGTCSAT GCTCGGGACA CGAACATGGG GCGGTGGTTT 540
 GAGGCGCAGG TGGTCAGGGT GACGCGGAAG GCCCCTCCCT GGGACGAGCC CTGCAGCTCC 600
 ACGTCCAGGC CGGCGCTGGA GGAGGACGTC ATTTACCAAG TGAATACGA CGACTACCCG 660
 GAGAACGCGG TGGTCCAGAT GAACTCCAGG GACGTCCGAG CGCGCGCCCG CACCATCATC 720
 AAGTGGCAGG ACCTGGAGGT GGGCCAGGTG GTCTGTCTCA ACTACAACCC CGACAACCCC 780
 70 AAGGAGCGGG GCTTCTGTTA CGACGCGGAG ATCTCCAGGA AGCGGAGAC CAGGACGGGG 840
 CGGGAACCTC ACGCCAACGT GGTGCTGGGG GATGATTCTC TGAACGACTG TCGGATCATC 900
 TTGTTGGACG AAGTCTTCAA GATTGAGCGG CCGGTTGAAG GGAGCCCAT GGTGACAAAC 960
 CCCATGAGAC GGAAGAGCGG GCGTCTCTGC AAGCACTGCA AGGACGACGT GAACAGACTC 1020
 TGCCGCTCTT GCGCTGCCCA CCTGTGCGGG GCGCGGCGAG ACCCGACAA GCAGCTCATG 1080
 75 TGCATGAGT GCGACATGTC CTTCACATC TACTGCTGG ACCCGCCCT CAGCAGTGT 1140
 CCCAGCGAGG ACGAGTGGTA CTGCCCTGAG TGCCGGAATG ATGCCAGCGA GGTGGTACTG 1200
 GCGGAGAGCG GGCTGAGAGA GAGCAAGAAG AAGGCGAAGA TGGCTCGGC CACATCGTCC 1260
 TCACAGCGGG ACTGGGCGAA GGGCATGGCC TGTGTGGGCC GCACCAAGGA ATGTACCATC 1320
 GTCCCGTCCA ACCACTACGG ACCCATCCCG GGGATCCCGG TGGGCACCAT GTGGCGGTTT 1380
 CGAGTCCAGG TCAGCGAGTC GGGTGTCCAT CGGCCCAAG TGGCTGGCAT ACACGGCCGG 1440
 80 AGCAACGAGG GAGGCTATCT CCGTGTCTG GCGGGGGGCT ATGAGGATGA CGTGAGCCAT 1500
 GGGAAATTTT TCACATACAC GGGTAGTGGT GGTGAGATC TTTCCGCAA CAAGAGGACC 1560
 GCGGAACAGT CTGTGATCA GAAACTCACC AACACCAACA GGGCGCTGGC TCTCACTGTC 1620
 TTTGCTCCCA TCAATGACCA AGAAGGGGCC GAGGCCAAGG ACTGGCGGTC GGGGAAGCCG 1680
 85 GTCAGGGTGG TGGCCTATGT CAAGGGTGGC AAGAATAGCA AGTACGCCCC CGCTGAGGGC 1740
 AACCGCTAGC ATGGCATCTA CAAGGTTTGG AAATACTGGC CCGAGAAGGG GAAGTCCGGG 1800
 TTTCTGCTGT GGGCTACCTT TCTGCGGAGG GACGATGATG AGCCTGGCCC TTGGACGAAG 1860
 GAGGGGAAGG ACCGATCRAA GAACTGGGG CTGACCATGC AGTATCCAGA AGGCTACCTG 1920

5
10
15
20
25
30
35

GAAGCCCTGG CCAACCGAGA GCGAGAGAAG GAGAACAGCA AGAGGGAGGA GGAGGAGCAG 1980
CAGGAGGGGG GCTTCGGGTC CCCCAGGACG GGCAGAGGCA AGTGGGAAGCG GAAGTCGGCA 2040
GGAGGTGGCG CGAGCAGGGC CGGGTCCCCG CGCCGACAT CCAAGAAAAC CAAGGTGGAG 2100
CCCTACAGTC TCACGGCCCC GCAGAGCAGC CTCATCAGAG AGGACAAGAG CAACGCCAAG 2160
CTGTGGAATG AGGTCCCTGGC GTCACTCAAG GACCGGGCCG CGAGCGGCAG CCGGTCCAG 2220
TTGTTCTCTG GTAAAGTGGA GGAGAOGTTC CAGTGTATCT GCTGTACGGA GCTGGTGTTC 2280
CGGCCCATCA CGACCGTGTG CCAGCACAAC GTGTGCAAGG ACTGCCTGGA CAGATCCTTT 2340
CGGGCAGCAGG TGTTTCAGCTG CCCTGCCTGC CGCTACGACC TGGGCGCGAG CTATGCCATG 2400
CAGGTGAACC AGCCTCTGCA GACCGTCTCT AACCAGCTCT TCCCGGCTA CGGCAATGGC 2460
CGGTGATCTC CAAGCACTTC TCGACAGGCG TTTTGTGTAA AACGTGTGCG AGGGCTCGTT 2520
CATCGGCACG GATTTTGTTC TTAGTGGGCT TAACTTAAAC AGGTAGTGT TCTCCGTTT 2580
CCTAAAAAGG TTTTCTCTCC TTTTCTTTTA TTTTCTTTT TCAATCTAT ACAATTTTTCAG 2640
GAATTTATGT ATTCTGGCTA AAAGTTGGAC TTCTCAGTAT TGTGTTAGT TCTTTGAAAA 2700
CATAAAGACC TGCAATTCTT CGACAAAACA ACACAAGATT TTTTAAAGAT GGAATCAGAA 2760
ACTACGTGTG GTGGAGGCTG TTGATGTTTC TGGTGTCAAG TTCTCAGAAG TTGCTGCCAC 2820
CACTCTTTA AGAAGGCGAC AGGATCAGTC CTTCTTAGG GTTCTGGCCC CCAAGGTGAG 2880
AGCAAGCATC TTCTGACAGC CATTTTGTCA TCTAAAGTCC AGTGACATGG TTTCCCGTGG 2940
TGGCCCGTGG CAGCCCGTGG CATGGCGTGG CTCAGCTGTC TGTGGAAGT GTTGCAAGGA 3000
AAAGAGGAAA GATCTCGGGC CTAGTTCAAA CCTTGGCTC AAAGCCATCC CCCACAGAC 3060
TGCTTAGCGT CTGAGATCCG GTGAAAAGT CCTCTGCCCA CGAGAGCAGG GAGTTGGGGC 3120
CACGCAGAAA TGGCCTCAAG GGGACTCTGC TCCAGTGGG GCCAGGCGTG GACTGACGC 3180
TGTCCGACGA AGGCGGCCAC GGAAGGACGC CAGCACAGCA AGTCAGTGC AAGTGCCTTT 3240
GATTGCTTCC TTCTTCTTAA AGACGACAGT CTTTGTGTT AGCACTGAAT TATTGAAAA 3300
GTCAACGCTA TTCTCAAAAC TGCAGTCAAT CAGTTCTTCC TGACACCGGA TGGGTGCTTG 3360
GGAACCGTTT GAGCCTTATA GATCATTTAC ATTCAATTT TTTAACTCAG CAAGTGAGAA 3420
CTTACAAGAG GGTTTTTTTT TAATTTTTTT TTCTCTTAAT GAACACATTT TCTAAATGAA 3480
TTTTTTTTGT AGTTACTGTA TATGTACCAA GAAAGATATA ACGTTAGGTT TTGGTTGTTT 3540
TTGTTTTTGT ATTTTTTTTC TTTTGAAAGG GTTGTGTAAT TTTTCTAATT TTACCAAGT 3600
TTGACGCTA TACCTCAATA AAACAGGGAT ATTTTAAATC ACATACCTGC AGACAAACTG 3660
GAGCAATGTT ATTTTAAAG GGTTTTTTTC ACCTCCTTAT TCTTAGATTA TTAATGTATT 3720
AGGGAAGAAT GAGACAATTT TGTGTAGGCT TTTTCTAAG TCCAGTACTT TGTCCAGATT 3780
TTAGATTCTC AGAATAAATG TTTTTCACAG ATTGAAAAA AAAAAAAA

Seq ID NO: 455 Protein sequence
Protein Accession #: NP_037414.2

40
45
50

1 11 21 31 41 51
MWIQVRTMDG RQTHTVDSLS RLTKVEELRR KIQELPHVEP GLQLFYRGK QMEDGHTLFD 60
YEVRLNDTIQ LLVRQSLVLP HSTKERDSEL SDTDSGCCIG QSESDKSSTH GEAAAEETDSR 120
PADEDMWDET ELGLYKVNRY VDARDTNMGA WFEAQVVRVT RKAPSRDEPC SSTSRPALEE 180
DVIYHVKYDD YPENGUVQMN SRDVRARART IIKWQDLEVG QVVMNLNPNP NPKERGFWYD 240
AEISKRRETR FARLEYANVV LGDDSLNDCR IIFVDEVFKI ERPGEGSPMV DNPMMRKS GP 300
SCKHCKDDVN RLCHVCACHL CGRQDDPKQ LMCDECDMAF HIYCLDPLLS SVPSDEWYIC 360
PECRNDASEV VLAGERLRES KKKAKMASAT SSSQRDWGKG MACVGRTEKC TIVPSNHYGP 420
IPGIPVGTMM RFRVQVSESG VHRPHVAGIH GRSNDGAYSL VLAGGYEDDV DHGNFFTYTG 480
SGGRDLSGNK RTAEQSCDQK LINTNRALAL NCFAPINDQE GAERKDWRSR KPVVRVVRNVK 540
GGKNSKYAPA BGNRYDGIYK VVKYWEKKG SGFLVWRYLL RRDDDEPGPW TKEGKDRICK 600
LGLTMQYPEG VLEALANRER EKENSKEEB EQQEGGFASP RTGKGKWKRK SAGGSPSRAG 660
SPRRTSKTKK VEPYSLTAQQ SSLIREDKSN AKLWNEVLAS LKDRPASGSP FQLFLSKVEE 720
TFQCICQCEL VFRPITTVQV HNVCKDCLDR SFRAQVPSCP ACRYDLGRSY AMQVWQLPQT

Seq ID NO: 456 DNA sequence
Nucleic Acid Accession #: NM_001200.1
Coding sequence: 325..1514

55
60
65
70
75
80
85

1 11 21 31 41 51
GGGGACTTCT TGAACCTGCA GGGAGAATAA CTTGCGCACC CCACTTTGCG CCGGTGCCTT 60
TGCCCCAGCG GAGCCTGCTT CGCCATCTCC GAGCCCCACC GCCCTCCAC TCCTCGGCTT 120
TGCCCCAGAC TGAGACGCTG TTCCACGCGT GAAAGAGAG ACTGCGCGGC CGGCACCCGG 180
GAGAAGGAGG AGGCAAGAA AAGGAACGGA CATTGCTTCC TTGCGCCAGG TCCTTTGACC 240
AGAGTTTTTC CATGTGGAAG CTCTTTCAAT GGAAGTGTCC CCGGTGCTT CTAGACGGA 300
CTGCGGTCTC TAAAGGTGCG ACCATGCTGG CCGGACCCCG CTGTCTTCTA CGCTTGCTGC 360
TTCCCCAGGT CCTCTCGGCG GCGCGGCTG GCCTGTTTCC GGAGCTGGGC CGCAGGAAGT 420
TGGCGGCGCG GTCTCGGGCG CGCCCTCAT CCCAGCCCTC TGAAGAGTGC CTGACGAGT 480
TGGAGTTGCG GCTGCTCAGC ATGTTGCGCC TGAACAGAG ACCCAACCCC AGCAGGGAAG 540
CGGTGCTGCC CCCCTACATG CTAGACCTGT ATCGCAGGCA CTCAGGTGAG CCGGGCTCAC 600
CCGCCCCAGA CCACCGGTG GAGAGGGCAG CCAGCCGAGC CAACACTGTG CGCAGCTTCC 660
ACCATGAAGA ATCTTTGGAA GAACTACCAG AAACGAGTGG GAAACAAACC CGGAGATTCT 720
TCTTTAATTT AAGTTCTATC CCCACGGAGG AGTTTATCAC CTCAGCAGAG CTTTCAAGTTT 780
TCCGAGAAAC GATGCAAGAT GCTTTAGGAA ACAATAGCAG TTTCCATCAC CGAATTAATA 840
TTTATGAAT CATAAAACCT GCAACAGCCA ACTCGAAATT CCGGTGACC AGACTTTTGG 900
ACACAGGTT GGTGAATCAG AATGCAAGCA GGTGGGAAAG TTTTGATGTC ACCCCCGCTG 960
TGATGCGGTG GACTGCACAG GGACACGCCA ACCATGGATT CGTGGTGGAA GTGGCCCACT 1020
TGGAGGAGAA ACAAGGTGTC TCCAGAGAC ATGTTAGGAT AAGCAGGTCT TTGACCAAG 1080
ATGAACAGAG CTGCTCAGC ATAAGGCCAT TGCTAGTAAC TTTTGGCCAT GATGGAAGA 1140
GGCATCTCTT CCACAAAGA GAAAAACGTC AAGCCAAACA CAACAGCGG AAACGCTTAA 1200
AGTCCAGCTG TAAGAGACAC CCTTTGTACG TGGACTTCAG TGACGTGGGG TGAATGACT 1260
GGATTGTGTC TCCCCCGGGG TATCACGCTT TTTACTGCCA CGGAGAAATG CCTTTTCTCT 1320
TGGCTGATCA TCTGAATCC ACTAATCATG CCAATGTTCA GACGTGGTGC AACTCTGTTA 1380
ACTCTAAGAT TCTTAAGCA TGCTGTGTC CGACAGAACT CAGTGTATC TCGATGCTGT 1440
ACCTTGACGA GAATGAAAAG GTTGTATTAA AGAATATCA GGACATGTTT GTGAGGGGTT 1500
GTGGGTGTCG CTAGTACAGC AAAATTAAAT ACATAAATAT ATATATA

Seq ID NO: 457 Protein sequence
Protein Accession #: NP_001191.1

1 11 21 31 41 51
 5 MVAGTRCLLA LLLPQVLLGG AAGLVPELGR RKFAAAASSGR PSSQPSDEVL SEFELRLLSM 60
 PGLKQRPPTS RDAVVPFYM DLVRRHSGQP GSPAPDHRLE RAASRANTVR SFHHEESLEE 120
 LPETSKITTR RFFFNLSISP TEEFITSDEL QVFREQMDA LGNNSSFHR INIYEIKPA 180
 TANSKFFVTR LLDT

Seq ID NO: 458 DNA sequence
 Nucleic Acid Accession #: NM_001999.2
 Coding sequence: 1..8736

1 11 21 31 41 51
 15 ATGGGAGAGAA GACGGAGGCT GTGTCTCCAG CTCTACTTCC TGTGGCTGGG CTGTGTGGTG 60
 CTCTGGGCGC AGGCGACGCG CGGCCAGCCT CAGCCTCCTC CGCCCAAGCC GCCCGGGCCC 120
 CAGCCGCGCG CGCAACAGGT TGGTTCGCT ACAGCAGGCT CTGAAGGCGG GTTCTTAGCG 180
 CCCGAGTATC GCGAGGAGGG TGCCGCACTG GCCAGCCGCG TCCGCGCGCG AGGACAGCAG 240
 GAGGTGCTCC GAGGGGCCAA CGTGTGCGGC TCCAGATTCC ACTCCTACTG CTGCCCTGGA 300
 20 TGGAAGACCG TCCTCGGAGG AAACAGGTGC ATTGTCCGA TTTGTAGAAA TAGTTGTGGA 360
 GATGGATTTT GTTCCCGTCC TAACATGTGT ACTTGTTCCT GTGGGCAAA ATCATCAACC 420
 TGTGGATGCA AATCAATTCA GCAGTGCAGT GTGAGATGCA TGAATGGTGG GACCTGTGCA 480
 GATGACCACT GCCAGTGCCA GAAAGGATAT ATTGGAACCT ATTGTGGACA ACCTGTCTGT 540
 GAAAATGGAT GTCAATATGG TGGACGTTGC ATCGCCCAAC CGTGTGCTTG TGTATTGGG 600
 25 TTCACTGGTC CACAGTGTGA AAGAGATTAC AGGACAGGCC CGTGTTCAC TCAGGTCAAC 660
 AACCAGATGT GCCAAGGGCA GCTGACAGGC ATTGTCTGCA CGAAGACTCT GTGCTGTGCC 720
 ACCACTGGAC GGGCGTGGGG CCATCCCTGT GAGATGTGTC CAGCCAGGCC TCAGCCCTGC 780
 CGACGGGGTT TCATCCCCAA CATCCGCACT GGAGCTTGCC AAGATGTTGA TGAATGCCAG 840
 GCTATCCAG GGTATGCTCA AGGAGGAAAC TGTATCAATA CAGTGGGCTC TTTTGAATGC 900
 30 AGATGCCCTG CTGTGCACAA ACAGAGTGAA ACTACTCAGA AATGTGAAGA CATTGTAGAG 960
 TGCAAGCATCA TTCTCGGGAT ATGTGAAACT GGTGAATGTT CCAACACCGT GGAAGCTAT 1020
 TTTTGTGTTT GTCCACGTGG ATATGTAAAC TCAACAGATG GCTCTCGATG CATGTCATCA 1080
 AGAACAGGCA TGTGTTCTCT GGGCCTGGTG AATGGCCGCT GTGCACAGA GCTCCCGGGG 1140
 AGAATGACGA AATGTCAGTG CTGCTGTGAG CTTGGCGCTG CTTGGGGCAT CGGAACCAT 1200
 35 CCTGAAGCCT GTCTCTGTC AGGTCTGTAG GAATATCCGA GACTTTGCAT GGATGGACTT 1260
 CCAATGGGAG GAATTCACAG GAGTGTGCTG TCCAGACCTG GAGGCACCTG GGGAAATGGC 1320
 TTTGCCCAAC GTGGCAATGG CAATGGCTAT GGGCCAGGAG GGACAGGCTT CATCCCATC 1380
 CCTGGAGGCA ATGGCTTTTC TCCTGGCGTT GGGGGAGCCG GTGTGGGGGC CGGGGGACAG 1440
 GGACCTTCCA TCACCTGGACT AACCAATTCTG AACAGACAA TAGATATCTG TAAGCATCAT 1500
 40 GCTAAACCTT GTTTAAATGG ACGCTGTATA CCAACTGTCT CAAGCTACCG ATGTGAATGC 1560
 AACATGGGTT ATAAGCAGGA TGCAAAATGGA GATTGTATAG ATGTTGATGA ATGCACATCA 1620
 AATCCCTGCA CTAATGGAGA TTGTGTTAAC ACACCTGTTT CCTATTATG TAAATGTGAT 1680
 GCTGGATTCC AGAGGACTCC TACCAAGCAA GCATGCATTG ATATTGATGA GTGCATCCAG 1740
 45 AATGGGGTTC TTTGTAAGAA CGTCTGATGC GTGAACCTAG ATGGAAGTTT CCAATGCATT 1800
 TGCAATGCCG GCTTTGAATT AACTACAGAT GGAAGAAACT GTGTTGATCA TGATGAATGT 1860
 ACAACTACCA ACATGTGTTT GAATGGAATG TGATCAATG AAGATGGCAG CTTCAAGTGC 1920
 ATCTGCAAAC CAGGATTGTT GTTGGCTCCA AATGGGCGTT ACTGTACTGA GTTGTATGAA 1980
 TGCCAGACCC CAGGAATCTG CATGAATGGG CACTGCATCA ACAGTGAAGG GTCCTTCGCG 2040
 50 TGTGACTGTC CCCGAGGCTT GGCCTGTGGC ATGGATGACG GTGTGTGTGT TGATCTCAC 2100
 ATGCGCAGTA CTTGCTATGG AGGAATCAAG AAGGAGTGT GTGTGGCTCC TTTCCCGGCT 2160
 GCAGTGACCA AGTCCGAATG CTGCTGTGCC AATCCAGACT ATGTTTGTGG AGAACCTTGC 2220
 CAGCCATGCA CTGCAAGAAA TTCACTGAAA TTCCACGGCC TTTGTAGTAG TGGAGTAGT 2280
 ATCACTGTGG ATGGAAGAGA TATCAATGAA TGTGCTTTGG ATCCTGATAT ATGTGCCAAT 2340
 GGGATTGTTG AAAAATTACG TGGTAGTTAC CGTTGTAATT GCAACAGTGG CTATGAACCA 2400
 55 GATGCTCTCG GAAGAACTGT TATTGACATT GATGAATGTT TAGTAAACAG ACTGCTTTGT 2460
 GATAACGGAT TGTGCCGAAA CAGCCAGGGA AGTTACAGCT GTACGTGCCC ACCAGGGTAT 2520
 GTGTTGAGGA CTGAGACAGA GACCTGTGAA GATATAAATG AATGTGAAG CAACCCATGT 2580
 GTCATGAGGG CCTGCAGAAA CAACCTTGGA TCTTTCAATT GTGAATGTTT GCCCGGCAGC 2640
 60 AAACCTCAGT CCACAGGATT GATCTGTATT GACAGCCTGA AGGGGACCTG TTGGCTCAAC 2700
 ATCCAGGACA GCCGCTGTGA GGTGAATATT AATGGAGCCA CTCTGAAATC TGAATGCTGT 2760
 GCCACCTCG GAGCCGCTGT GGGGAGCCCC TGTGAGCGGT GTGAACCTGA TACAGCTTGC 2820
 CCAAGAGGGC TTGCCAGGAT TAAAGGTGTT ACGTGTGAAG ATGTTAATGA GTGTGAGGTG 2880
 TTCCCTGGCG TTTGTCCAAA TGGACGCTGT GTCAACAGTA AGGGATCTTT TCATTGCCAG 2940
 65 TGCCCTGAAG GCCTTACGTT GGATGGGACT GGGCGTGTAT GTTTGGATAT TGCCATGGAG 3000
 CAGTGTACT TGAAGTGGGA TGAAGATGAA TGCAATCCACC CGTTCCTGGA AAAGTTCCGC 3060
 ATGGATGCCT GCTGCTGTGC TGTGCGGGCG GCTTGGGGCA CGAGTGTGTA GGAGTGGCCC 3120
 AAACCTGGCA CCAAGGAATA CGAGACTG TGCCCCCGCG GGGCTGGCTT TGCTAACCGA 3180
 GGGGATGTTT TTAATGGGCG GCCATTTTAC AAAGACATCA ATGAATGCAA AGCATTTCCT 3240
 70 GGGATGTGCA CTTATGGGAA GTGCAGAAAT ACAAATCGAA GCTTCAAATG CGTTGCAAT 3300
 AGTGGCTTTG CTCTAGACAT GGAGGAAAGA AACTGCACGG ACATCGACGA GTGCAGGATT 3360
 TCTCTGACC TCTGTGGCAG TGAATCTGCG GTCAATACAC CGGGCAGCTT TGAGTGGCAG 3420
 TGCTTCGAAG GCTATGAAG TGGCTTCATG ATGATGAAGA ACTGCATGGA CATTGAOOGA 3480
 TGTGAACGTA ACCCTCTCCT TTGTAGGGGT GGCACCTGTG TGAACACTGA GGGCAGCTTT 3540
 75 CAGTGTGACT GCCCACTGGG ACACGAGCTG TCACCATCCC GTGAGGACTG TGTGATATT 3600
 AATGAATGCT CCCCTGAGTG CAATCTCTGC AGAAATGGAA AATGTGTGAA CATGATTGGA 3660
 ACCTATCAGT GCTCTGTCAA TCCTGGATAT CAGGCTACGC CAGACCGCCA GGGCTGTACA 3720
 GATATTGATG AATGTATGAT AATGAACGGA GGCTGTGACA CCCAGTGAC ACATTCAGAG 3780
 GGAAGCTAGC AATGCAGCTG CAGTGGGGT TATGCCCTGA TGCCAGATGG GAGATCGTGT 3840
 GCAGACATGA ATGAATGTGA AAACAATCCT GATATCTGTG ATGGCGGCCA GTGTACCAAC 3900
 80 ATTCTGGAG AGTATCGCTG CCTCTGCTAT GATGGCTTCA TGGCTTCAT GGACATGAAA 3960
 ACATGCAATG ATGTCAATGA ATGTGACCTA AATTCAAATA TCTGCATGTT TGGGGAATGT 4020
 GAGAACACAA AGGGATCCCT CATTTGCCAC TGTCACTGG GTTACTCAST GAAGAAGGGG 4080
 ACCACAGGAT GTACAGATGT GGAATGGTGT GAAATGGTGT CTCATAACTG CGACATGCAT 4140
 GCCTCATGTC TGAATATCCC AGGAAGCTTC AAGTGTAGCT GCAGAGAAGG CTGGATTGGA 4200
 85 AACGGCATCA AGTGTATTGA TCTGGACGAA TGTCTAATG GAACCCACCA GTGTAGCATC 4260
 AATGCTCAGT GTGTAAATAC CCCGGGCTCA TACCGCTGTG CTGTCTCGA AGGTTTCACT 4320
 GGTGATGGCT TTACCTGCTC AGATGTTGAT GAGTGTGCGA AAAACATAAA CCTCTGTGAG 4380

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

AACGGACAGT GCCTTAATGT CCGGGGTGCA TATCGCTGCG AGTGTGAGAT GGGCTTCACT 4440
CCAGCCTCAG CCAGCAGATC CTGCCAAGAT ATTGATGAAT GCTCCTTCCA AAACATTGT 4500
GTCTCTGGAA CATGTAATAA CCTGCCCTGA ATGTTTCATT GCATCTGCGA TGATGGTTAT 4560
GAATTGGACA GAACAGGAGG GAACCTGTACA GATATTGATG AGTGTGCGA TCCTATAAAC 4620
TGTGTCAATG GCCTATGTGT CAACACGCCT GGTGCTGATG AGTGTAACTG CCCACCCGAT 4680
TTTCAGTTGA ACCCAACTGG TGTGGGTTGT GTTGACAACC GTGTGGGCAA CTGCTACCTG 4740
AAGTTTGGAG CTGAGGAGGA TGGGAGTCTG TCTTGCAACA CCGAGATCGG GGTGGGGGTC 4800
AGTGCCTCTT CATGCTGCTG CTCTCTGGGA AAGGCCCTGG GAAACCCCTG TGAGACATGC 4860
CCCCCTGTCA ATAGCACTGA ATATTACACC CTGTGTCCTG GAGGTGAAGG CTTGAGACCT 4920
AACCCCATCA CAATCATTTT AGAAGACATT GACGAATGCC AGGAGTTACC AGGTCTCTGC 4980
CAGGGTGGAA ACTGCATCAA CACTTTTGGG AGCTTCCAGT GTGAGTGCCC ACAAGGCTAC 5040
TACCTCAGCG AGGATACCCG CATCTGTGAG GATATTGATG AGTGTTTTGC ACATCCTGGT 5100
GTGTGTGGGC CTGGGACCTG CTATAACACC CTGGGAAAT ACACCTGCAT TTGCCCACT 5160
GAGTACATGC AGGTCAATGG AGGCCACAAC TGCAATGGACA TGAGAAAAG CTTTGTCTAC 5220
CGAAGCTATA ATGGAACAC TGTGTGAGAAT GAGTTGCCCT TCAATGTGAC AAAAAGGATG 5280
TGCTGTGCGA CATATAATGT GGGCAAAGCT GGGAAACAAC CTGTGAAACC ATGCCCAACT 5340
CCAGGAACAG CTGACTTTAA AACCATATGT GGAAATATTC CTGGATTAC CTTTGACATT 5400
CACACAGGAA AAGCTGTGA CATGTATGAA TGTAAAGAGA TTCCAGGCAT TTGTGCAAT 5460
GGTGTGTGCA TTAACAGAT TGGCAGTTTC CGCTGTGAAT GCCCTACAGG ATTCAGTTAC 5520
AATGACCTGC TGTGTGTTTG TGAAGATATA GATGAGTGCA GCAATGGTGA TAATCTCTGC 5580
CAGCGGAATG CAGACTGCAT CAATAGTCCCT GGTAGTTACC GCTGTGAATG TGCCCGGGGT 5640
TTCAAATTT CACCAATGG GGCCTGTGTA GATGCGAATG AATGTTTGA AATTCCTAAC 5700
GTTTGCAGTC ATGGCTTGTG TGTGTATCTG CAAGGAAGTT ACCAGTGCAT CTGCCCAAT 5760
GGCTTTAAGG CTCTCAGGA CCAGACCATG TGCATGGATG TTGATGAGTG CGAGCGCAC 5820
CCATGTGGAA ATGGAACCTG TAAAAACACC GTTGGATCCT ATAACGTCT CTGCTACCCA 5880
GGGTTTGAAC TCACTCATAA TAAATGATTG CTGACATAG ATGAGTGACG TTCTTTTTT 5940
GGTCAGGTGT GCAGAAATGG ACGTTGTTTT AATGAAATG GTTCTTTCAA GTGCTATGT 6000
AACGAAGGTT ATGAACCTAC CCCAGATGGC AAAAACTGTA TAGACACTAA TGAGTGTGTC 6060
GCCCTTCCCG GCTCTTGCTC TCCCTGTACC TGTCAGAAAT TGGAGGGATC CTTGAGATGC 6120
ATCTGTCCCC AAGCGGATGA AGTAAAAAGC GAGAACTGCA TTGATATAAA TGAATGTGAT 6180
GAAGATCCCA ACATTGTCTT TTTTGGTTCC TGTAATAA CTCCAGGGGG CTTCCAGTGC 6240
CTCTGCCCCC CTGCTTTTGT ACTATCTGAT AATGGACGGA GATGCTTTGA TACTGCCAG 6300
AGCTTCTGCT TCACAAATTT TGAATAAGGA AAGTGTCTG TACCAAAGC TTTCAACACC 6360
ACAAAAGCAA AATGTCTGCTG TAGTAAGATG CCAGGAGAGG GCTGGGGGGA CCCCTGTGAG 6420
CTGTGCCCCA AAGCAGATGA AGTTGCATT CAGGATTTGT GTCCATATGG CCATGGAAC 6480
GTCCCTAGTC TTATGATAC ACGTGAAGAT GTCAATGAGT GTCTTGAGAG CCCAGGCATT 6540
TGTTCAAATG GTCATATGAT CAACACCGAC GGATCTTTTC GCTGTGAATG TCCATGGGC 6600
TACAACCTTG ACTACACTGG AGTACGCTGT GTGATACTG ATGAGTGTTC AATCGGCAAT 6660
CCGTGTGGAA ATGGTACATG CACCAATGTT ATTGGGAGTT TTGAATGCAA TTGCAATGAA 6720
GGCTTTGAGC CAGGGCCCAT GATGAATTGT GAAGATATCA ACGAATGTGC CCAGAACCCA 6780
CTGCTGTGTG CTTTACGCTG CATGAACACT TTTGGGTCCT ATGAATGCAC GTGCCCGATT 6840
GGCTATGCCC TCAGGGAAGT TCAAAAGATG TGCAAGATC TGGATGAATG TGCTGAAGGG 6900
TTACAGACT GTGAATCTAG GGGCATGATG TGTAAGAATC TAATCGGCAC CTTTATGTGC 6960
ATCTGCCCTC CTGGAATGGC CCGAAGGCCG GATGGAGAAG GCTGTGTAGA TGAATAAGAA 7020
TGCAAGACCA AGCGGCAAT CTGTGAAAT GAGAGTTGTG TTAACATTAT TGAAGCTAT 7080
AGATGTGAGT GATATGAAGG ATTCAGTCA AGTCTCTCAG GCACTGAATG CCTTGACAA 7140
CGACAGGGTC TCTGCTTTCG AGAGGTACTG CAGACAATAT GTCAATGGC ATCCAGTAGT 7200
CGCAATCTCG TCACTAAGTC AGAATGCTGC TGTGATGGTG GGCAGGGCTG GGGCCACCAG 7260
TGGCGACTTT GCCCACTTCC TGGAACTGCC CAGTACAAAA AGATATGTCC CTATGGCCCA 7320
GGATATACAA CTGATGGAAG AGATATTGAT GAATGTAAAG TAATGCCAAA CCTCTGACC 7380
AATGTTCAGT GCATCAATAC CATGGGCTCA TTCCGATGCT TCTGCAAGGT TGGCTACACC 7440
ACAGACATCA GTGGAACCTC TTGTATAGAC CTTGATGAAT GCTCCAGTC CCGAAACCA 7500
TGCAACTACA TCTGCAAGAA CACTGAGGGG AGTTATCAGT GTTCAATGCC GAGGGGGTAT 7560
GTCTCTCAAG AGGATGGAAA GACATGCAAA GACCTTGATG AATGTCAAAC AAAGCAGCAT 7620
AATCTGCCAT TCTCTGTGT CAACACCCCTG GGGGGGTTTA CCGTAAATG TCCACTGGT 7680
TTCAACAGAG ATCACTCTGC TTGTATCGAC AACACCAAT GTGGGTCTCA ACCTTTGCTT 7740
TGTGGAGGAA AGGGAATCTG TCAAAACACT CCAGGCAGTT TCAGCTGTGA ATGCCAAAGA 7800
GGGTTCCTCT TTGATGCCAC CGGACTGAAC TGTGAAGATG TTGATGAATG TGATGGGAAC 7860
CACAGGTGCC AACACGGCTG CCAGAACATC CTGGGTGGCT ACAGATGTGG CTGCCCCCAA 7920
GGCTACATCC AGCACTACCA GTGGAATCAG TGTGTGATG AGAATGAATG CTCCAATCCC 7980
AATGCTGTG GCTCTGCTTC CTGCTACAC ACCCTGGGGA GTTACAAGTG CGCTGCCCC 8040
TGGGGTTCT CTTTCAGACA GTTCTCCAGT GCCTGCCACG ACGTGAATGA GTGCTGCTCC 8100
TCCAAGAAC CCGCAATTA CGGCTGCTCT AACACGGAGG GGGGCTACCT CTGTGGCTGC 8160
CCCCCTGGT ATTACAGAT GGGACAAGGC CACTGTGTCT CAGGAATGGG ATTTAAACAAG 8220
GGGCAGTACC GTTCACTGGA TACAGAGGTC GATGAGGAAA ATGCTCTGTC CCCAGAAGCA 8280
TGCTACAGAT GCAAAATCAA CGGCTATCCT AAGAAAGACA GCAGGCAGAA GAGAAGTATT 8340
CATGAACCTG ATCCCACTGC TGTGAAACAG ATCAGCCTAG AGAGTGTGCA CATGGACAGC 8400
CCCGTCAACA TGAAGTTCAA CCTCTCCAC CTGGCTCTA AGGAGCACAT CTGGGAAC 8460
AGGCCGCCCA TCCAGCCCT CAACAACCA ATCCGTTATG TCATCTCTCA AGGGAACGAT 8520
GACAGCGTCT TCCGATCCA CCAAGGAAT GGGCTCAGCT ACTTGACAC GGGCAAGAA 8580
AAGCTCATGC COGGACATA CACACTGGAA ATCACTAGCA TCCTCTCTA CAAGAAGAAG 8640
GAGCTTAAGA AACTGGAAGA GAGCAATGAG GATGACTACC TCCTAGGGGA GCTTGGGGAG 8700
GCTCTCAGAA TGAGGCTGCA GATTACAGTC TATTAAACCG TCACAGACTT GGGCCAGGC 8760
TCAATCTCTA GCACAGCCAG TCTGCAAGG CATTGAAAA GTCAAGGACT AATTTTAAAG 8820
AGGAAAAATA ATAATAACTC TTGTTTCTTT CCTCCCTGTC TTAGACTTTG AATGTTGACC 8880
CTCACAGGGA GGGATAATTT AGACTCTGGT ATGGCCAAAG ATTTGAGCTC AAAGGCAACC 8940
GTGGTTACTG TATTTTTTAT ATAACCTCAT TTTAAATAT ATTTAAAGAA ACCTAAATGT 9000
TCAAGATATC AGCATATGGC ACTAAATGCA CAAAAATAAT GTGAGCTTTT TTTTTTTTT 9060
CCTGTAGCA GTCTGTAACT CTTTGGGTAT TTTGCTATAG TTGCTAATTA AAAAAATATA 9120
GATGTTTAT TATTTTTAAT GCAGTAATAT ATGGAGAAAT GAACAACTA TGTAAACAAA 9180
AAGGAAACT CACTTGTTTT TCTTTAGATT TATAAATTTG AGCTATTTTT TTTAGAGGTG 9240
CTTTTAAAA ATCCAATAGA TACAAGAGAT GTTTCCTTTG GTTTCTGCC AGTCAATCCAG 9300
CTGATACACA CCTGATGAT TTTAAAGAAA GCCACACAGA GCTGAATCGG GCAAGTCTAA 9360
TCAATATTT AAAAGACATG AATGTCTATTA GATCCTTTAT AAGTAGATC GAAGCCAAAG 9420
CAGCTCATTT GTGACACAT TTCATATCAC CAGACACACC AGGCAACAGA AGTTGAAGCA 9480
CAACCACTGT AGCAAAATAC CTTGACTGCT TGTGAGACCA TTAGCATTGC AGGCCAAACC 9540
GTACTGTATT TCCTTCTCAT AACCTCAAGG AACCATATGT GTTACCCACA ACACCTCAT 9600

WO 02/086443

PCT/US02/12476

5
10
15
20
25
30
35
40
45
50
55
60
65

```

CTTACCCAGG GTGCGCTGCG TCCTCATGGT ACTGTAGGCA GCTGAAGAAC CGCCGTTCOC 9660
TTGAAAGGGA ACACCTGGCA TTCTGTGGTG TTCTGTGCTG TCTTAAATAA TGGTGCAITTT 9720
ATTATGTTCA AGTATTATTC GGATTGCCAT ATGTGCAAAAC AAATCATGCA ATGCAGCCAA 9780
DGAATATATG TTGTTGTTGT TGTITTAAC CCATTTTTTT TTAGAATTTT TCATTAATAC 9840
TGATAGTATA CACCATATGC CTCATTTTAT CATAGCCCTAT TGTGTATGAA AGATGTTTGT 9900
ACAAAGAAAT GATGTTTAGT TTGCTTTAGT CATTAAAAAA GATATTGTAC CAGGATGTGC 9960
TATTAAGAGC ACGTATCCAT TATTCTTCTC AACCCAAGAA CTGTTTCCTT GGACCAAGTA 10020
CCAAACCTCA TATGTGAAAT GGCCAAAGCA CATGCAGGCT CTGGTTGTTT CCTCTCAAA 10080
CTGTGCTGAC CAAAGATTAG TAACCAAGTA TACCCAGTAT TTTGAGGTTT TATTGTTTTT 10140
TAAATAACTA AAAAAAACT CGTGCC

```

Seq ID NO: 459 Protein sequence
Protein Accession #: NP_001990.1

15
20
25
30
35
40
45
50
55
60
65

```

1 11 21 31 41 51
MGRRRRLCLQ LYFLWLGCVV LWAQGTAGQP QPPPKPRPRP QPPQPVRSR TAGSEGGFLA 60
PEYREEGAIV ASRVRRRQQQ DVLRGPNVCG SRPHSYCCPG WKTLPGGNQC IVPICRNSCG 120
DGFCSRPNMC TCSSGQISST QGSKSIQQCS VRCMNGGTCA DDHCQCGQGY IGTYCGQPV 180
ENGCGNGGRG IAPQPCAVYG FTGPQECERDY RTGPGCTQVN NMQCQCGQLTG IVCTKTLCCA 240
TTGRANGHPC BMCFAQPPCP RRGFIPNIRT GACQDVDECO AIPGICQGGN CINTVGSFEC 300
RCPAGHKQSE TTQKCEDIDE CSIIPGICET GECSNTVGSY PCVCPRGYVT STDGSRCIDQ 360
RTGMCPSGLV NGRCAQLPBG RMTKMQCCCE PGRCWGIGTI PEACFVRGSE EYRRLCMDGL 420
PMGIPGSAG SRPGGTGGNG FAPSGNGNGY GPGGTGFPII PGNGFSPGV GGAGVGAGGQ 480
GPIITGLTIL NQITIDICKH ANLCLNGRCI PTVSSYRCEC NMGYKQDANG DCIDVDECTS 540
NPCTNGDCVN TFGSYCKCH AGFQRTPTKQ ACIDIDEICQ NGVLCNGRC VNSDGSFQCI 600
CNAGFELTDD GKNVVDHDEC TTTNMCLNGM CINEDGSFKC ICKPGFVLAP NGRYCTDVDE 660
CQTPGICMNG HCINSEGSFR CDCPPGLAVG MDGRVCVDTH MRSTCYGGIK KGVCVRPFP 720
AVTKSECCCA NPDYGFGEPC QPCPAJNSAE FHGLCSSGVG ITVDGRDINE CALDPDICAN 780
GICENLRGSY RCNCSNGYEP DASGRNCIDI DECLVNRLLC DNGLCRNTPG SYSTCPPGY 840
VFRTEETECB DINECESNPC VNGACRNLG SFNCECSPGS KLSSTGLICI DSLKGTCLN 900
IQDSRCEVNI NGATLKSECC ATLGAANGSP CERCELDIAC PRGLARIKGV TCEDVNECEV 960
PPGVCNPRGK VNSKSGFHCE CPEGLTLDGT GRVCLDIRME QCYLKWDEDE CIHPVPGKFR 1020
MDACCAVAGA NPGTCECEPC KPGTKEYETL CPRGAGPANR GDVLTGRFFY KDINECKAFP 1080
GMCTYKCRN TIGSFKCRN SGFALDMEER NCTDIDECRI SPDLCGSGIC VNTPGSFCE 1140
CFEGYESGFM MMKNCMDIDG CERNPLLCRG GTCVNTGESP QDCPLGHEL SPSREDVDI 1200
NECSLSDNLC RKGKCVNMIG TYQSCNPGY QATPDROGCT DIDECEMIMG GCDTQCTNSE 1260
GSYECSCSEB YALMPDGRSC ADIDECENNP DICDGGQCTN IPGEYRCLCY DGFMAISDMK 1320
TCIDVNECDL NSNICMFEPC ENTGSGFICH CQLGYSVKKG TTGCTDVDEC EIGAHCNDMH 1380
ASCLNTPGSF KCSCREGWIG NGIKCIDLDE CSNGTHQCSI NAQCVNTPGS YRCACSEGFT 1440
GDGPTCSDVD ECAENINLCE NGQCLNVPGA YRCECEMGFT PASDSRSCQD IDECSFQNIC 1500
VSGTCNINLP MFHICDDGY ELDRITGNCT DIDECAPIIN CVNGLCVNTP GRYECNCPD 1560
PQLNPTGVGC VDNKVGNCYL KFGPRGDSGL SCNTEIGVGV SRSSCCSLG KAWGNPCETC 1620
PPVNSTEYIT LCPGEGEPRF NPTIILEDI DECQELPLGL QGGNCINTFG SFQCECPQGY 1680
YLSIEDTRIC DIDECAHPHG VCGPOTCYNT LGNYTCICPP EYMQVNGGHN CDMRKSFCY 1740
RSYNGTTCEN ELPPNVTKRM CCCTYNVGKA GNKPCEPCPT PGTADEFKTI C GNIPGFTFDI 1800
HTGKAVDIDE CKZIPGICAN GVCINQIGSF RCECPTGFSY NDLLLVCEDI DECSNGDNLC 1860
QRNADICINP GSYRCECAAG FKLSFNGACV DRNECLEIPN VCSHGLVDL QGSYQICHN 1920
GFKASQDQTM QMDVDECERH PCGNGTCKNT VGSYNCLCYP GFELTHNNDL LDIDECSFF 1980
GQVCRNGRCF NEIGSFKCLC NEGYELTPDG KNCIDTNECV ALPGSCSPGT CQNLEGSFRC 2040
ICPPGYEYKS ENCIDINECD EDPNICLFGS CTNTPGGFQC LCPPGFVLSD NGRRCFDRQ 2100
SFCTPNEFG KCSVPKAFNT TKAKCCSKM PEGGWGDPCE LCPKDEVAF QDLCPYGHGT 2160
VPSLHDTRED VNECLESPGI CSNGQCINTD GSFRCECPMG YNLDTGVRC VDTDECSIGN 2220
PCNGTCTNV IGSFECNCE GFEPGPMNC EDINECAQNP LLALRCMNT FGSYCTCPI 2280
GYALREDQKM CKDLDECAEG LHDCESRGM CKNLIGTFMC ICPPGMARRP DGBGCVDENE 2340
CRTPKPGICN GRVNIIGSY RCECNEGFQS SSSGTECLDN RQGLCPAEVL QTICQMASS 2400
RNLVTSECC CDGGRGWGHQ CELCPLPSTA QYKICPHGP GTTIDGRID ECKVMPNLCT 2460
NGQCINTMGS FRCPCKVGT TDISGTSCID LDECSQSPKP CNYICKNTEG SYQCSCPRGY 2520
VLQEDQKTK CKDQCKTKH NCQFLCVNTL GGFTCKCPFG FTQHTACID NNECQSPL 2580
CGGKICQNT PGSPSCBQR GFSLDATGLN CEDVDECDGN HRCQHGQNI LGGYRCGPQ 2640
GYIQHWQWQ CVDNECSNP NACGSASCYN TLGSYKACAP SGFSFDQPS ACHDVNECS 2700
SKNFCNYGS NTBSGYLCG PPGYRVGQG HCVSGMGFNK GQYLSLDETV DEENALSPEA 2760
CYECKINGYP KDSRQKRSI HEPDPTAVEQ ISLESVDMDS PVMKFNLSH LGSKEHILEL 2820
RPAIQPLNKH IRYVISQND DSVFRIHORN GLSYLHTAKK KLMPTVYLE ITSIPLYKKK 2880
ELKKLEESNE DDYLLGELGE ALMRLLQIQ Y

```

Seq ID NO: 460 DNA sequence
Nucleic Acid Accession #: NM_013372.1
Coding sequence: 63..617

70
75
80
85

```

1 11 21 31 41 51
GCGGCGCAC TCAGCGCCAC GCGTCGAAAG CGCAGGCCCC GAGGACCGCG CGCACTGACA 60
GTATGAGCGG CACAGCCTAC ACGGTGGGAG CCTGCTTCT CCTCTGGGG ACCCTGCTGC 120
CGGCTGCTGA AGGGAAGGAG AAAGGTGCCA CCCCCGCCA GACAAGGCC 180
AGCACATGA CTCAGAGCAG ACTCAGTCGC CCCAGCAGC TGGCTCCAGG AACCGGGGG 240
GGGCGCAAGG GCGGGCACT GCCATGCCCC GGGAGGAGGT GCTGGAGTCC AGCCAGAGG 300
CCCTGCATGT GACGAGCGC AAATACCTGA AGCGAGACTG GTGCAAAACC CAGCCGCTTA 360
AGCAGACCAT CCACGAGGAA GGCTGCAACA GTCCGACCAT CATCAACCGC TTCTGTTAG 420
GCCAGTGCAA CTCCTTTTAC ATCCCCAGGC ACATCCGAA GGAGGAAGGT TCCTTTAGT 480
CCTGCTCCTT CTGCAAGCCC AAGAAATTC ATACCATGAT GGTCACTC AACTGCCCTG 540
AACTACAGCC ACCTACCAAG AGAAGAGAG TCACACGTGT GAAGCAGTGT CGTTGCATAT 600
CCATCGATTG GGATTAGGCC AAATCCAGGT GCACCCAGCA TGTCTAGGA ATGCAGCCCC 660
AGGAAGTCCC AGACCTAAAA CAACCAGATT CTTACTTGGC TTAACCTAG AGGCCAGAG 720
AACCCTCAGC TGCTCTCTG CAGGAGCCTG CTTGTGCGTA GTTCGTGTG ATGAGTGTG 780
ATGGGTGCTT GTGGGTGTTT TTAGACACCA GAGAAAAAC AGTCTCTGCT AGAGAGCACT 840
CCCTATTTTG TAAACATATC TGCTTTAATG GGGATGTACC AGAAACCCAC CTCACCCCG 900

```

5

10

15

20

25

30

35

40

45

50

55

CTCACATCTA AAGGGGCGGG GCGTGGTCT GGTCTGACT TTGTGTTTT GTGCCCTCCT 960
GGGGACCAGA ATCTCCTTTC GGAATGAATG TTCATGGAAG AGGCTCCTCT GAGGGCAAGA 1020
GACCTGTTTT AGTGCTGCAT TCGACATGGA AAGTCTCTTT TAACCTGTGC TTGCATCCTC 1080
CTTTCTCCTC CTCTCTCACA ATCCATCTCT TCTTAAGTTG ATAGTGACTA TGTCACTCTA 1140
ATCTCTTGTG TGCCAAAGTT CCTAAATTAA TTCACITTAAC CATGATGCAA ATGTTTTTCA 1200
TTTTGTGAAG ACCCTCCAGA CTCTGGGAGA GCGTGGTGTG GGCAGGACA AGCAGGATAG 1260
TGGAGTGAGA AAGGGAGGGT GAGGGGTGAG GCCAARTCAG GTCCAGCAAA AGTCAGTAGG 1320
GACATTGACG AAGCTTGAAA GGCCAATACC AGAACACAGG CTGATGCTTC TGAGAAAGTC 1380
TTTTCTAGT ATTTAACAGA ACCCAAGTGA ACAGAGGAGA AATGAGATTG CCAGAAAGTG 1440
ATTAACCTTG GCGTGTGCAA TCTGCTCAA CCTAACACCA AACTGAAAAC ATAAATACTG 1500
ACCACCTCTA TGTTCGGACC CAAGCAAGTT AGCTAAACCA AACCAACTCC TCTGCTTTGT 1560
CCCTCAGGTG GAAAAGAGAG GTAGTTTAGA ACTCTCTGCA TAGGGGTGGG AATTAATCAA 1620
AAACCKCAGA GGCTGAAATT CCTAATACCT TTCTTTTATC GTGGTTATAG TCAGCTCATT 1680
TCCATTCCAC TATTTCCCAT AATGCTTCTG AGAGCCAATA ACTGATTGA TAAAGATCCT 1740
GCCTCTGCTG AGTGTACCTG ACAGTAAGTC TAAAGATGAR AGAGTTTAGG GACTACTCTG 1800
TTTTAGCAAG ARATATTTTG GGGTCTTTT TGTTTTAACT ATTGTGAGGA GATTGGGCTA 1860
RAGAGAGAGC GACAGAGATA AGGAAATAAA GGGRATTGCC TCTGGCTAGA GAGTAAGTTA 1920
GGTGTTAATA CCTGTAGAAA ATGTAAGGGA TATGACCTCC CTCTCTTAT GTGCTCACTG 1980
AGGATCTGAG GGGACCCCTG TAGGAGAGCA TAGCATCATG ATGTAITAGC TGTTCATCTG 2040
CTACTGGTGT GATGGACATA ACTATTGTAA CTATTCACTA TTTACTGGTA GGCACCTGCC 2100
TCTGATTAAT CTTCGGCTAC TGGCAATGGC TACTTAGGAT TGATCTAAGG GCCAAAGTGC 2160
AGGGTGGGTG AACTTTATTG TACTTTGGAT TTGGTTAACC TGTTTTCTTC AAGCCTGAGG 2220
TTTTATATAC AAACCTCCCTG AATACTCTTT TTGCTTTGTA TCTTCTCAGC CTCTAGCCA 2280
AGTCTATGT AATATGSAAA ACAACACTG CAGACTTGAG ATTCACTTGC CGATCAAGGC 2340
TCTGGCATTG AGAGAACCTT TGCAACTCGA GAAGCTGTTT TTTTTCGTT TTTGTTTTGA 2400
TCCAGTGCTC TCCCATCTAA CAATAAACA GGAGCCATT CAAGGGCGGA GATATTTTAA 2460
ACACCCAAAG TGTGGGCTC GATTTTCAAA CTTTTAACT CACTACTGAT GATTCCTCAG 2520
CTAGGCGAAT TTGTCCAAAC ACATAGTGTG TGTGTTTTGT ATACACTGTA TGACCCCAAC 2580
CCAAATCTTT GTATTGTCCA CATCTCCAA CAATAAAGCA CAGAGTGGAT TTAATTAAGC 2640
ACACAAATGT TAAGGCAGAA TTTTGAGGGT GGGAGAGAAG AAAAGGGAAA GAAGCTGAAA 2700
ATGTAAACAC ACACAGGGA GGAATAATGA CATTGAGAAC CAGCAACAC TGAATTTCTC 2760
TTGTGTTTTT AACTCTGCCA CAAGAATGCA ATTTCTGTTA TGGAGATGAC TTAAGTTGGC 2820
AGCAGTAATC TTCTTTTAGG AGCTTGTACC ACAGTCTTGC ACATAAGTGC AGATTGGGT 2880
CAAGTAAAGA GAATTTCTC AACACTAATC TACTGGGAT AATCAGCAGC GTAACCTACC 2940
TAAAGCATA TCACTAGCCA AAGAGGAAA TATCTGTTCT TCTTACTGTG CCTATATTAA 3000
GACTAGTACA AATGTGGTGT GTCTTCCAAC TTTCAITGAA AATGCCATAT CTATACCATA 3060
TTTTATTGGA GTCACTGATG ATGTAATGAT ATATTTTTC ATTATTATAG TAGAATATT 3120
TTATGCAAG ATATTGTGG TCTGATCAT ACCTATTAAA ATAATGCCAA ACACCAAATA 3180
TGAATTTTAT GATGTACACT TTGTGCTTGG CATTAAGA AAAAAACCA CATCTGSA 3240
GTCTGAAGT TGTGTTTTGT TACTGTAGGT CTCAAGTT AAGAGTGTAA GTGAAAAATC 3300
TGGAGGAGAG GATAATTTCC ACTGTGTGGA ATGTGAATAG TTAATGAAA AGTTATGGTT 3360
ATTAAATGTA ATATTACTT CAATCCTTT GGTCACTGTG ATTTCAAGCA TGTTTCTTT 3420
TTCTCCTTTA TATGACTTTC TCTGAGTTGG GCAAAGAAGA AGCTGACACA CCGTATGTTG 3480
TTAGAGTCTT TTATCTGGTG AGGGGAAACA AAATCTTGAC CCAGCTGAAC ATGCTCTCCT 3540
GAGTCAGTGC CTGAATCTTT ATTTTAA TGTGAATGTC CTAAAGGTT AACATTTCTA 3600
AAGCAATATT AAGAAAGATC TTAATGTTA TTTTGAAGA CTTACGATGC ATGTATACAA 3660
ACGAATAGCA GATAATGATG ACTAGTTCAC ACATAAAGTC CTTTAAAGGA GAAATCTAA 3720
AATGAAAGT GGATAAACAG AACATTATA AGTGATCAGT TAATGCCATA GAGTGAAAGT 3780
AGTTCTATTG ACATTCCTCA AGATATTATA TATCAACTGC ATTAGTATAT ATGTCTGCTT 3840
AAATCATTTA AAAACGGCAA AGAATTATAT AGACTATGAG GTACCTTGCT GTGTAGGAGG 3900
ATGAAGGGG AGTTGATAGT CTCATAAAC TAATTGGCT TCAAGTTTCA TGAATCTGTA 3960
ACTGAATTT AATTTTACC CCAATAATGT TCTATATAGC CTTTGCTAAA GAGCAACTAA 4020
TAAATTAAC CTATCTTTT AAAAAAAA

Seq ID NO: 461 Protein sequence
Protein Accession #: NP_037504.1

60

1 11 21 31 41 51
MSRTAYTVGA LLLLLLTLFP AAEKKKKGSQ GAIPPPDKAQ HNDSEQTQSP QPGSRNRGR 60
GQGRGTAMPG EEVLESSQEA LHVTERKYLK RDWCKTQPLK QTIHEEGCNS RTIINRPCYG 120
QCNSFYIPRH IRKEBSPQS CSPCKPKKPT TMWVTLNCPB LQPPTKKKRV TRVKQCRCIS 180
IDLD

65

Seq ID NO: 462 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..2733

70

75

80

85

1 11 21 31 41 51
ATGAAAGTTG GAGTGTCTGT GCTCAITTTCT TTCTTCACCT TCACTGACGG CCACGGTGGC 60
TTCTCTGGGA AAAATGATGG CATCAARACA AAAAAAGAAC TCATTGTGAA TAAGAAAAAA 120
CATCTAGGCC CAGTCGAAGA ATATCAGCTG CTGCTTCAGG TGACCTATAG AGATTCCAAG 180
GAGAAAAGAG ATTGTAGAAA TTTTCTGAAG CTCTTGAAGC CTCCATTATT ATGGTCACAT 240
GGGCTAATTA GAATTAATCAG AGCAAAGGCT ACCACAGACT CCAACAGCCT GAATGGAGTC 300
CTGCAGTGTA CCTGTGAAGA CAGCTACACC TGGTTTCTCT CCTCATGCCT TGATCCCCAG 360
AACTGCTACC TTCACAAGGC TGGAGCACTC CCAAGCTGTG AATGTCTCTT CAACAACCTC 420
AGCCAGAGTG TCAATTTCTG TGAGAGAACA AAGATTGGG GCACCTTTCAA AATTAATGAA 480
AGGTTTACAA ATGACCTTTT GAATTCATCT TCTGCTATAT ACTCCAAATA TGCAAAATGGA 540
ATTGAAATTC AACTTAAAAA AGCATATGAA AGAATTCAAG GTTTTGAATC GGTTCAGGTC 600
ACCCAATTTT GAAATGGAAG CATCGTTGCT GGGTATGAAG TTGTTGGCTC CAGCAGTGCA 660
TCTGAACCTG TGTACGCCAT TGAACATGTT GCCGAGAAGG CTAAGACAGC CCTTCACAAG 720
CTGTTTCCAT TAGAAGACGG CTCTTTCAGA GTGTTCCGAA AAGCCAGTG TAATGACATT 780
GTCTTTGGAT TTGGGTCCAA GGATGATGAA TATACCTTGC CTGCAAGCAG TGGCTACAGG 840
GGAACATCA CAGCCAAAGT TGAGTCTCT GGGTGGCAGG TCATCAGGGA GACTTGTGTG 900
CTCTCTCTGC TTGAAGAAGT GAACAAGAAT TTCAGTATGA TTGTAGCAA TGCCACTGAG 960
GCAGCTGTGT CATCTCTCGT GCAAAATCTT TCTGTCTATCA TTGGGCAAAA CCCATCAACC 1020

	ACAGTGGGGA	ATCTGGCTTC	GSTGGTGTG	ATTCTGAGCA	ATATTTCATC	TCTGTCACTG	1080
	GCCAGCCATT	TCAGGGGTGC	CAATTCAACA	ATGGAGGATG	TCATCAGTAT	AGCTGACAA	1140
	ATCCTTAATT	CAGCCTCAGT	AACCAACTGG	ACAGTCTTAC	TGCGGGAAGA	AAAGTATGCC	1200
5	AGCTCAGCGT	TACTAGAGAC	ATTAGAAAAC	ATCAGCACTC	TGGTGCCTCC	GACAGCTCTT	1260
	CCTCTGAATT	TTTCTCGGAA	ATTCAATTGAC	TGGAAAGGGA	TTCCAGTGAA	CAAAAGCCAA	1320
	CTCAAAAGGG	GTTACAGCTA	TCAGATTAAA	ATGTGTCCCC	AAAATACATC	TATTCCTCAT	1380
	AGAGGCCGTG	TGTTAATTGG	GTCCAGACCA	TTCCAGAGAT	CCCTTCCAGA	AACTATTATC	1440
	AGCATGGCCT	CGTTGACTCT	GGGGAACATT	CTACCCGTTT	CCAAAAATGG	AAATGCTCAG	1500
10	GTCAATGGAC	CTGTGATATC	CACGGTTATT	CAAAACTATT	CCATAAATGA	AGTTTTCCTA	1560
	TTTTTTTCCA	AGATAGAGTC	AAACCTGAGC	CAGCCTCATT	GTGTGTTTGT	GGATTTTCAGT	1620
	CAITTCGAGT	GGAACGATGC	AGGCTGCCAC	CTAGTGAATG	AAACTCAAGA	CATCGTGACG	1680
	TGCCAATGTA	CTCAGTTGAC	CTCCTTCTCC	ATATTGATGT	CACCTTTTGT	CCCCTCTACA	1740
	ATCTTCCCGC	TGTGAAAATG	GATCACCTAT	GTGGGACTGG	GTATCTCCAT	TGGAAGTCTC	1800
	ATTTTATGGC	TGATCATCGA	GGCTTTGTTT	TGGAAGCAGA	TTAAAAAAG	CCAAACCTCT	1860
15	CACACACGTC	GTATTGTCAT	GGTGAACATA	GCCTGTGCCC	TCTTGATGTC	TGATGTCTGG	1920
	TTTATTGTTG	GTGCCACAGT	GGACACCAAG	GTGAACCTCT	CTGGAGTCTG	CACAGCTGCT	1980
	GTGTTCTTTA	CACACTTCTT	CTACCTCTCT	TGTGTTCTCT	GGATGCTCAT	GCTTGGCATC	2040
	CTGCTGGCTT	ACCGATTCAT	CCTCGTGTTC	CATCACATGG	CCCAGCATTT	GATGATGGCT	2100
20	GTGGGATTTT	GCCTGGGTTA	TGGGTGCCCT	CTCATTTAT	CTGTCAATAC	CATTGCTGTC	2160
	ACGCAACCTA	GCAATACCTA	CAAAAGGAAA	GATGTGTGTT	GGCTTAACTG	GTCCAAATGGA	2220
	AGCAAAACCAC	TCTTGCTTTT	TGTTGTCCCT	GCCTGGCTTA	TGTTGGCTGT	GAACTTCGTT	2280
	GTGGTGTCTG	TAGTCTCTAC	AAAGCTCTGG	AGGCGGACTG	TTGGGGAAAG	ACTGAGTGGG	2340
	GATGACAAGG	CCACCATTCAT	CCGCGTGGGG	AAGAGCCTCC	TCATTCTGAC	CCCTCTGCTA	2400
25	GGGCTCACCT	GGGCTTTTGG	AATAGGAACA	ATAGTGGACA	GCCAGAATCT	GGCTTGGCAT	2460
	GTTATTTTTG	CTTTACTCAA	TGCATTCCAG	GGATTTTTTA	TCTTATGCTT	TGGAATACCT	2520
	TTGGACAGTA	AGCTGGCACA	ACTTCTGTTT	AACAAGTTGT	CTGCCCTAAG	TTCTTGGGAG	2580
	CAACACAGAA	AGCAAAACTC	ATCAGATTTA	TCTGCCAAAC	CCAAATCTCT	AAAGCCTTTC	2640
	AACCCACTGC	AAAACAAAGG	CCATTATGCA	TTTTCTCATA	CTGGAGATTC	CTCCGACAA	2700
30	ATCATGCTAA	CTCAGTTTGT	CTCAATGAA	TAA			

Seq ID NO: 463 Protein sequence
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
35	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
40	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
45	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
50	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	

Seq ID NO: 464 DNA sequence
Nucleic Acid Accession #: AB035089.1
Coding sequence: 9845..10219

	1	11	21	31	41	51	
60	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
65	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
70	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
75	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
80	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
85	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	
	1	11	21	31	41	51	

	TGGAGAGAG	TCTGGCATT	CCTCAAAATG	TTAACTCGGA	TTTACCATAT	GACCCAGGSA	1800
	TTTCAATCAT	AGGTTTATAC	TCAAAAGAAA	TGAAGAAATA	TGCCATGCAA	AAAAATGTAC	1860
	ATGAAAGGTC	ACAAATCAT	TATTCATAAT	AGTAAAGGGA	TGGAACCAAC	ACAAATGTCC	1920
5	ATCAACTTAT	GATTAAAGAA	AATCTGGTCT	ATTCTAGAAA	TGGAATATTA	TTGCACCCCA	1980
	AAAAGGAATG	ATGTACTGAT	CCATGCAATG	ATGTGGACAA	ACCATGAAAA	TAACACTAGA	2040
	TTAAAGAAAG	CAGTCACAAA	AGGACTTACT	GTATGATTCC	ATTTACCTGA	AATGTTTGA	2100
	ATAGGCAAA	CCATAGAAAC	AGGAGGTAGA	TTCTGGTTT	CCAGGTCTC	CAGGAAGGGA	2160
	AGAATGAAGT	ACAAGATTTC	TTTGGAGGT	AGTGAATTTG	TGTGGAATG	AGATCATGAT	2220
10	GATGATAGCA	CAACTTTGTG	AATATAATAA	AATCATTGAA	TGTACAGTT	GAATTTATGG	2280
	TATATAAATT	ATATGTTAAT	AAAAAGGGGG	TCCACAAAAC	AAACAGCCCC	CACTCTGGT	2340
	TGTCAGGGAG	ATATTGGATT	AAATGGCCTT	GGACAAACAC	CCCTCTCCCT	GGCCACAGAC	2400
	ATTCTTCAGA	TTACAAGATA	TTCCAGGGGA	AACACTGGAA	TGAGTCTGAA	GCCAGTGTCT	2460
	AAACAGAAAG	ACCATTGAGA	AATGTTGTGA	TCTTGACAGG	TCAAGCAATT	TATTTTTCGG	2520
	CTTCATTTT	AAATGTAAAA	TTAGAAAGCT	GCCATTTAAA	ATGGCCCGTC	TGTTTCAATT	2580
15	GCTCTTCTCA	GTGTCAGCCT	GTAACTCAA	TGTGTTAGTC	TGTTTTCATG	CTGCTGATAA	2640
	AAACATACTC	GAGACTGGCA	AGAAAAAGAG	GTTTAAATGG	GCTTAGAGTT	CCACGTGATT	2700
	GGGGAGGCGT	CAGAAATACA	GTAGGAGGCA	AAAGTTATTC	TTACATGGTG	GCTGCAAGAG	2760
	AAGATGAGGA	AGAAGCAAAA	GAAGAAACCC	CTGATAAAC	CATCGGATCT	CCTGAGGCTT	2820
	ATTAACATATC	ATGAGAATAG	CACAAGAAAG	ACCGCCCCCT	ATGATTCAAT	TACCTCTACC	2880
20	TGGGTCCCTC	CAATAACATG	TGGAATTTCT	GGTAGATACA	ATTCAAGTTG	AGATTTGGGT	2940
	GGGAACACAG	CCAAACCATTA	TCACTCAGCA	AGGCAGATAA	CTTTCTCACT	GAGCCTATGC	3000
	AACAGAAAAC	CATCTGGGAT	GGTTGTAAGG	GGCAGAGGAA	GTGACTGGTA	GGATCACTGC	3060
	CAAAGCTGAG	CACCTCAGAG	AAGGCAATAG	AATCTTATTC	TCCATAGTAT	GCTATAAGAT	3120
	ACTGAAGTAC	ACTTCTTCTC	TATCTCTTTG	GACTTAGAAT	TAGCACTACA	TTCTTTGTTA	3180
25	TACAGAAAAA	TACTAAGGA	AATTCATAGG	ATGACAAAAA	CTTTCAGAAC	TGAAAAACAG	3240
	GAATATGAAG	CTTTTATAGT	CTTTGGTATT	CGAAGTATGC	CTAAAAGACA	ATGCAAAATC	3300
	CAAGAAAGAA	ATGGTGGGGT	TTTTGTTTGT	TTGGTTTTGT	TTTTGTTTTA	CAGCTGGAGT	3360
	AGAATACAAA	GGGATGGAGT	TGAAACAAAT	GAGAGGAAAT	TGGAATTTCT	AATTTATTCT	3420
	CATTGGCATT	AGAAAGGCAC	CTACATGTAT	TTACATGAG	CCGGTGACTG	CTGACTTGCA	3480
30	TTCTTATTTT	TTCCCTATAG	ATTAAAAAGG	AGGTACAATG	GTAGAACTGT	AATCCTGTCC	3540
	TTTGTCATAA	ATTTTCATAT	TCAATAAGGT	GAGTGTATGC	CCGCTTGTA	AATCTGAAGT	3600
	TGAGTAACTT	CAATTAATAA	CCACAGAGGG	AAAGGCAGCA	AGAGGAGAGG	CATAAATTTA	3660
	GGATCTCACC	CTTCATTCCA	CAGACACACA	CAGCCTCTCT	GCCCACCTCT	GCTTCTCTTA	3720
	GGAAACAGAG	TAGAGCTTTC	AAGCCTCTCC	AGCTTAATAA	CATGAATTAT	TTTTGAGAAT	3780
35	AATAATGATA	CTGTGTTCTA	TATCATGCAT	CTCTGCTATT	CTGTCTGATT	ATATTTTACT	3840
	TATTCCTCCA	GAGCAAAAT	AAAAATACCTA	TTTCATCTGA	TTTGCTCTTT	ATCTAAATTTG	3900
	CTTAGTTCCA	AGTAAACCAA	GGCACTTTTA	GGAAACAGCA	GGGAGAGTGC	CTTGAGGCCA	3960
	GAGAGTCTTG	AAGGAGATGT	CAGGGACGCA	TCTTAACAGC	TGGTTGGATG	TGATCCACAG	4020
	AGGTCTCTCG	TTAGCATTTA	TTGTAAGGCC	ATCCTACCTA	GCTCTAGTGT	AACCAGCAAT	4080
40	GAAGAGAAAG	TAAAGAGGGT	CGATTACTTA	TTTACAATAG	TCTTTAAAAA	CGTAGTTTTG	4140
	TAAAGCTTCT	AATTAGGACA	TTAATATATT	TAATATATGC	ACATTGTAGA	AAGATTGAAG	4200
	CGTTAAAAAT	AAGAGAAAAA	CTTTAAATGT	CAAAATCTCA	CAACCCAGAT	ATATCATTTT	4260
	TTTTAAGAAA	TTGTACTACA	AAATACCATT	CCATTTATTA	AAGTCATTCT	GACAGGAATC	4320
	TGATGCTTTT	CCAGGAGTTC	CAGATCACAT	CGAGTTCACT	ATGAATTCAC	TCAGTGAAGC	4380
45	CAACACCAAG	TTTCATGTTG	ATCTGTTCCA	ACAGTTCCAG	AAATCAAAAG	AGAACAAACAT	4440
	CTTCTATTCC	CTATTCAGCA	TCACATCAGC	ATTAGGGATG	GTCCTCTTAG	GAGCCAAAGA	4500
	CAACACTGCA	CAACAAATTA	GCAAGGTAGC	TATCAGCATC	ATTAAGTTGT	CTGTGTCAG	4560
	TTTTTCTCTG	GTTCGGTGGG	CTAGCACGCA	GATGGTAATA	GATGTGGTGG	TCTGATGGGT	4620
	AGCACAGGGG	GCTGTGACAG	AATTTCCATA	ACTGTGAGAC	CAGTGACTTA	AACAGATCTT	4680
50	TTGAGTAAAG	TTTTCTTGTC	CGCTTTCATG	TCTCTTCCAG	GTTCTTCACT	TTGATCAAGT	4740
	CACAGAGAAC	ACCACAGAAA	AAGCTGCAAC	ATATCATGTG	AGTCACAGAG	CACCTGTGAT	4800
	CAGCTTTAGA	TCCCTGAAAC	GGTCATAGTT	TAAACCTGGA	ACTTCACAAA	AACATAAGAA	4860
	AGGCCAGTTT	TAGGGAAAAAT	CTTGACACCA	AAGATTGAGA	CATACAGAGT	GGGTTGGCAT	4920
	TTCATGGCAC	ATAATTATTA	TTCTCTATT	CTGCGTTACT	AAAGACAGT	CAGCACTGTA	4980
55	CCTCAGAGAT	TAGGCTGGA	TCAGGATAGG	CTGGGTTTCA	ACTCCAGCTT	TGCTCTTCA	5040
	AAATGATGAA	TAAAGCAGG	ACACAATGCT	TOGGAGTCCC	AGTGACCTCA	TCCAGAAAAA	5100
	CTAAGGGTAA	GAATAAATCT	GACTCAATAC	ATGCAAAATG	TTACAACAGT	TTAATATATC	5160
	GCTTGGCCCA	TAAAGTGCAT	AATAAATGTT	ATTATTATTA	TAAAGTAGCT	ATAATTATAC	5220
	TAATCATAAAT	AATGTGAAAA	TAATTTAATT	TTCAATGAGT	CATTAAATGAG	ATTGAGAGGA	5280
60	ATAAGCACAA	GTCCTAAGTAT	ATTTTGGAAA	ATGATTGCTA	TGGAATATAT	TGGTTTAGAG	5340
	CCTTAATAGT	GCATAAATGCT	TTGCTGGAAG	GTAGAAAGTT	CTAGATTTAA	ACAGGCTTAG	5400
	GTTCAAAAT	TGGCACTTCT	AATTTATGTC	TCTATAAACA	GGGTTTTTTT	CCCCATTCTC	5460
	TGAGCTTTCT	TGTGTTTCAT	TGAATTGAAC	TAAAGACTTA	GAGTTACCCA	TGTAAGTGCC	5520
	TTAGCCATGG	ACCTGGCATA	CACCTCTCTT	ACGTGACAGG	AATGACCATC	ATGAGGAAAG	5580
65	AGCCACAGAT	CAGTCAATGT	GTCTTACAG	ATAATAGCAC	CAACAGGTAT	AACAGGGCTT	5640
	CCTGGCATAA	TCTATTTAAA	ATATCCAAAC	TTCAACATAC	TGATATCCTT	GATGACTGTT	5700
	AGAAGTGAAA	TATGGTCCCT	GCCCATAGG	AGCTGAGAGT	TTAACTGGGA	AGCTAAACCT	5760
	AACCCCTTAA	ACCAACAAGG	AGAAAACTTA	CTGGTAGACA	GCGCTGCATC	TTAGTTTCAG	5820
	AAGAGAAAAA	ATTGCAGTAC	GTTAGAGCAA	GAAGAAATTT	CTGGAAGAA	TCAATATATA	5880
70	GGTGGATTTT	GAAGGTGATT	TGAGGTGAAA	TACACCAATT	ATCAGGGAAT	AACATCAAGG	5940
	GTCTCTCAAT	AGACTACAG	CATTAGGGA	CTGATCTAAC	AGACTTAGCA	TGGGTTTAGT	6000
	ATTTACATTG	ATACAGCAAT	TGAATGATCT	CCTTTTTTGA	TGTTTGAAGG	TTGATAGGTC	6060
	AGGAAATGTT	CATCACCGAT	TTCAAAAGCT	CTGACTGAA	TTCAACAAAT	CCACTGATGC	6120
	ATATGAGCTG	AAGATCGCCA	ACAAGCTCTT	CGGAGAAAAA	ACGTATCAAT	TTTACAGGT	6180
75	AATTTCACTT	GGCTTACCCA	CATTTCAATT	GCATCTGAT	GTCTGTGCTC	CTGAGTGGCC	6240
	AAATGGAAGA	AAGCAAGGCA	GATGAGCCTG	GCGAGCCCA	GTGGAGAGCA	TTTACTCAGA	6300
	GTGCATTAGC	TCCATTTCOA	CAACTCTCCC	CCACTGGAGT	GTCCAGACC	CCAACGATAC	6360
	ATCATCTGAAG	TGTGGATTTA	GGGATAATCT	TGTGATAAAA	GAGGAGGTTG	TGTAATAGAG	6420
	TGAGTAAAGG	TAAATAGTAA	TAAAGATACCA	TGATAAAACT	GGCACTGACT	CAGTCACTA	6480
80	CGATACATCT	TGTTGGGAAA	TGTATGACTA	ATGGGATATT	ATTGGAATGG	CGAGGCTTGG	6540
	GTGAGTCTCT	GAGAATAGTT	GAGGAAGTAC	CAGGAATAT	TGAATGCACA	GGATGAAAGA	6600
	CAAAACACAA	GATCAGAAAC	ATCATGGTTA	AAATTAAGTG	AGAGAAGTCT	GAGAAGCAAT	6660
	GAATCTCTCT	CAGGGAAGCC	TGCTCTGACG	TTTGCAAAAC	ACAGCCTCTT	CTGCTTCTGC	6720
	CTTTTGCCAA	GATGATATTG	ACCTTCAGTG	ACCTCTTCTT	TGTGCCAGCC	CACATTCCCC	6780
85	TTTTGCACTG	CCTCATGAC	ACCTGTATAA	AAATATCCAT	GGACAGGAGA	TACTGCATCT	6840
	ATTCAGGGTC	TGGATTGAGC	TTACTGTTGT	TACAAATAAG	TAAAGTTGGT	AATATATAGT	6900
	TACATAAATT	ACTCTTAATT	CCTACTTCTT	CCTTCATATC	TCAAGGAAT	ATTAGATGTC	6960

	CATCAAGAAA	TTTTACCAGA	CCAGTGTGGA	ATCTACTGAT	TTTGCAAATG	CTCCAGAAGA	7020
	AAGTCGAAAG	AAGATTAAC	CCTGGGTGGA	AAGTCAAACG	AATGGTAGGA	GAGCCACCCA	7080
	TTATAGAAAC	ACCTTTGAGA	AACCTATGCC	AGTGAGCCTT	GTGCTTGACA	CTGCATGGGG	7140
5	GAACAGGTGT	GGGGAATGAG	ATGGGTTTGC	AGGGAGGGCT	GAAGAGGGCA	CTCCAGATGA	7200
	AGGATTTGTC	CAATGAATA	TGAAGAGAGC	CTAGGGGAGC	CAAGGAGGAA	ATCACAGGAA	7260
	GCCAAATTAGA	TGGAACACA	TCTGGAGAAT	TATTTGCTTA	TGGCCCTGCA	TGACATATGC	7320
	TTTGTGGATC	CCCTGTCTCC	GCTCAGACCT	ATTTTGAGAT	CATATCCTTT	ACTTTAAATC	7380
	AGACTCAAA	TTTTATGATG	AATATTTAAT	AGAAAACATT	AGAAAGCGTC	TCTCGTCTCC	7440
10	TTTACTAAT	GGGAAACAAG	CAGCTCTCTG	GTAAATCAAC	CTTTTGTCTC	TGAGCTGGAG	7500
	CTGCCTGGAT	CACATCTGTA	GCCATGTGT	TCTGCAGGGA	TTATCACAGC	TCTCTTCCCC	7560
	ATCAAGGGCA	AAGAGCTTGA	CAAAGTCTCC	ATTCTACAGA	CATCTTTCTT	ACCTCCCAAC	7620
	TCTCATTACA	GGCCAACTTT	ACAGCAACTC	AACATGAGAG	TGAATAGGAA	GATACCCCCG	7680
	GAAGTAGTGT	CTGACAGCAC	AGGACATGCG	TTTCATATTA	CAGAGCTCAA	GTCACTCATC	7740
15	CTAAAAATG	AGCAGGGCCT	CCTTCTCTG	AATGGGGACC	CGTAGTTAA	AAAAAAATAA	7800
	AAGTAGGAAG	AGGAGGGAGG	GAGAAAGGAA	AGACACATGT	TGGAAGAGTA	GACAAAATCA	7860
	GTTTATCAGT	ATTCCAAAT	AGATGATTGG	AGACATTCT	ACACAGAGAA	CGTGAACCTC	7920
	TTCTCTATCA	CAAGAAGTGA	TGTCTCCATC	AAGGGTAACT	TTATACGACT	GGAGCCTTGA	7980
	AGAAAGCTGC	ATCTGGTGAA	CCACTGTGTA	GTGAGTCTAA	CAATTCAAAG	ATCAAAGTCA	8040
20	GTGAGTCTCA	AGCAGGGTAA	TGGGTCAATA	ATTAACGATC	AGTCACGAAC	ATTGCAAAAG	8100
	CATCTTCCAG	ACAAGCCATT	TGTAGCTTGT	GTAAAGACT	CTTTTATTCT	TTCCCTTGCA	8160
	GAAAAAATTA	AAAACCTATT	TCTGTATGGG	ACTATTGGCA	ATGATACGAC	ACTGGTTCTT	8220
	GTGAACGCAA	TCTATTTCAA	AGGGCAGTGG	GAGAATAAAT	TTAAAAAAGA	AAACACTAAA	8280
	GAGGAAAAAT	TTGGCCAAA	CAAGGTATTG	TCTATATTTT	ATTTATATAG	TGTAATATGT	8340
25	TAATACATGG	AATGTTAAAC	ATTTCTGATG	GAATGTAACA	TGATAAGTAA	AAAAATAAAA	8400
	TTGTTTATGT	CTGTTATTTT	GTTGTTTTAC	TCTTATAACT	TTATTTAGTT	AGGAATACCT	8460
	GAAAAACTAT	TGTTTCTAAC	TCATGGAATT	CCTGGGTTAT	TTCTTAGAAG	AAGAAGGATG	8520
	TGTTGCTATC	TCAATAATAT	TATCTTTTTT	GTCTTGTGTT	TCACGTGTTA	TTTGTGGAC	8580
	ACATTGATT	ATTGCAGAA	ACATACAAAT	CTGTACAGAT	GATGAGGCAA	TACAATTOCT	8640
30	TTAATTTTGC	CTTCTGGAG	GATGTACAGG	CCAAGTCTC	GGAAATACCA	TACAAAGGCA	8700
	AAGATCTAAG	CATGATTGTG	CTGCTGCCAA	ATGAAATCGA	TGGTCTGCAG	AAGSTAAGAA	8760
	CTTGATCTTA	CAACTCTTCC	TTCTACTGCC	GGACATTTTT	CCAAAGATAC	CAAGTTTAAA	8820
	CAAGGTAAAC	GCTTATGACC	GAGTTGCCTC	AAAATGATGA	AAAATTCTAA	ATGAGGAATG	8880
	ATGACTCACC	TTCATATTAC	AAATATTGTA	GCATAGGGCC	TGACACAAAC	TGAAGCTTAA	8940
35	GTTTTGTTT	GTTTGTGTT	TTTTATTATT	ATTATTATAA	TACTTTAAGC	TTTAGGGTAC	9000
	ATGTGCACAA	TGTGCGAGTT	AGTTACATAT	GTATACATGT	GCCATGCTGG	TGTGCTGCAC	9060
	CCATTAACTC	ATCATTTAGC	GTTAGGTATA	TCTCCTAATG	CTATCCCTCC	CCCTCTCCCC	9120
	CACCCCAACA	CAGTCTCCAC	AGTGTGATGT	TACCTTCTGT	TGTCCAAGTG	TTCTCATGTT	9180
	TCAATTCCCA	TCTATGATT	AATCCATCT	ATGGCTTAGT	TAATGATTAA	TTTATTAGAG	9240
40	TTACATGCAT	TGATATCAA	TTTGATGATA	TTATTATGCA	GCAATTTAAA	CTTGACTGGG	9300
	AGAAATATAT	ACCAATGTGA	GGAAAGTTTA	CAAATAGGCC	GAGTAGAAAA	GGGAATACAA	9360
	ATTTAGGAAT	TTAGGGAATT	ACAATTTAAT	AATTGCAATG	TGTACTAAAT	AATGTATACA	9420
	GAAAAATATG	ATGAGCCTAT	TAAAAATTGA	CACATGTAGT	AGGCTGTTGG	CACAAGAAAT	9480
	AGTGATACAT	ACAGTTCATT	GTGTACAAAA	TAATGTAATC	ATATTTTACA	TGTGTATCAT	9540
45	ACAGTTGTAT	ACATACATAT	GTACACATAT	ACATATACGT	AAAAACATGA	TTCTGTTTTT	9600
	ACATACATGT	ATATACATAT	ACACATATAA	CCCAATGTAT	TTATATATTC	AGGACTCATA	9660
	TTTTACCTAT	TAGAATAATA	ATGCTCTATTA	AAGTGAACCT	TCTGTATTTC	ACATTTATITG	9720
	CCAAAATAAC	GAATCTCCAC	ATAGTCAATT	CATTGTAAAG	GTGTATTAGA	GATCGACAGT	9780
	TAGTCATATC	AGTTTCTTTT	TTCCATTGTT	ATAGCTTGAA	GAGAAACTCA	CTGCTGAGAA	9840
50	ATTGATGGAA	TGGCAAGATT	TGCAGAAAT	GAGAGAGACA	TGTGTCGATT	TACACTTACC	9900
	TGGTTTCAAA	TGGTAAGAGT	GCTATGACCT	CRAGGACACG	TTGAGAACCA	TGGGAATGGT	9960
	GAATATCTTC	AATGGGAGTG	CAGACCTCTC	AGGCATGACC	TGGAGCCACG	GTCTCTCAGT	10020
	ATCTAAAGAT	CTACACAAGG	CCTTTGTGGA	GGTCACTGAG	GAGGGAGTGG	AAGCTGCAGC	10080
	TGCCACCGCT	GTAGTAGTAG	TGGAATTATC	ATCTCCTTCA	ACTAATGAAG	AGTTCGTGTG	10140
55	TAATCACCTT	TTCTATTCT	TCATAAGGCA	AAATAAGACC	AACAGCATCC	TCTTCTATGG	10200
	CAGATTCTCA	TCCCATAGTA	TGCAATTAGT	CTGTCACTCC	ATTTAGAAAA	TGTTCACTTA	10260
	GAGGTGTTCT	GGTAAACTGA	TTGCTGGCAA	CAACAGATTG	TCTTGGCTCA	TATTTCTTTT	10320
	CTATCTCATC	TGTATGATGA	TAGTCATCAT	CAAGAATTTA	ATGATTAAAA	TAGCATGCCT	10380
	TTCTCTCTTT	CTCTTAATAA	GCCACATAT	AAATGTACTT	TTCTTCCAG	AAAAATTTCC	10440
60	CTTGAGGAAA	AATGTCCAAG	ATAAGATGAA	TCATTTAATA	CCGTGTCTTC	TAAATTTGAA	10500
	ATATAATTCT	TGTTCTGACC	TGTTTTAAAT	GAACCAAAAC	AAATCATACT	TTCTCTTCAA	10560
	ATTTAGCAAC	CTAGAAACAC	ACATTCTTTT	GAATTTAGGT	GATACCTAAA	TCCTTCTTAT	10620
	GTTTCTAAAT	TTTGTGATTC	TATAAAACAC	ATCATCAATA	AAATAATGAC	ATAAATCAT	10680
	TTTTGCTTTA	CCTGTTTTCT	CTCTGAAAAG	GGCAAGTGTC	CAGTTACACA	TAGGAAAGAT	10740
65	AATTTAGAGA	TATATTAATC	ATATATAAAG	GAAAAATAAA	AACAGAGTAG	TTCATGATGA	10800
	GCCTGGAGTA	GAAGGCATAT	CCCAGAACAG	GAGGAGCCTT	GTAACCCACA	TAGGAACCTC	10860
	CTATTTTATG	CTAAGGGATG	AAGAAACTCA	TTACAGGCTT	TGATGGTTGT	TTGTCAAAGA	10920
	GGGGCATAAG	ATTATCATAT	CCACATCTAG	AAAAATACATC	TCTGGCTACG	CTGATATCAA	10980
	TGGATGCGAG	GAAAGAACAG	TGTGTTTACC	ATATATAAAT	TAGGAAATCA	TTAGAGTATT	11040
70	GGGAGTGAAA	ATGGAGAGAA	AGAAAGAGCC	TGGGGGAATT	ATTTAGGAAA	TAATAGTTAC	11100
	AGAAAGACAT	CTAAGTTGCT	GACCTATCTG	ACTGGATGGA	TGGAAGAATA	TCTTGTCTCT	11160
	GAGAGAAAAA	AAGACTTTGG	GTTTAAATTT	GTACTTGTATG	AAATTAAGGTA	CTTTTAATAT	11220
	TCAATGGAT	TTGCCCTGGCA	GGCACTTGAA	GATATTAGTC	TAAATCTCAG	AAACAGAATA	11280
	TGATCTGAAG	CTCTAAATTT	GTGATATTCA	ATATAAATAC	TTTAGAGTCA	TTGGGATAAA	11340
75	TATGGTAGTT	GTAGCTAAAA	GCAAAAAATAA	GATACCTAGG	AGAAAGGATA	AAGTTAGAAG	11400
	AAAGAAGAAT	CTAGAATTGA	CCTTGAAGTA	TATCAGCATG	TGTAAGATC	AGGAATTGAT	11460
	CATTTTATT	TTCCAGAAAG	TAGCTTTTCT	TAGGGTTCCA	TATTTACTCC	CATAGATTCT	11520
	TCCC						

Seq ID NO: 465 Protein sequence
Protein Accession #: BAB21525.1

1	11	21	31	41	51	
MNSLSEANTK	FMPDLFQQFR	KSKENNIFYS	PISITTSALGM	VLLGAKDNIA	QQISKVLHFD	60
QVTEHTTEKA	ATYHVDKSGN	VHQFQKLLT	EFNKSTDAYE	LKIANKLFGK	KTYQPLQEYL	120
DAIKKPYQTS	VESTDFANAP	EESRKKINSW	VESQTNKIK	NLFPDGTIGN	DTTLVLVNAI	180
YFKQWENKF	KKENTKEEKF	WPNKNTYKSV	QMMRQNSFN	FALEDVQAK	VLEIPYKGD	240

LSMIVLLPNE IDGLQKLEEK LTAELMEWT SLQNMRETCV DLHLPRFKME ESYDLKDTLR 300
TMGMVNIENG DADLSGMTWS HGLSVSKVLH KAFVEVTEEG VEAAAATAVV VVELSSPSTN 360
EEPCNHPLF PFIRQNKNTS ILFYGRFSSP

5 Seq ID NO: 466 DNA sequence
Nucleic Acid Accession #: NM_001910.1
Coding sequence: 50..1240

10 1 11 21 31 41 51
GGAGAGAAGA AAGGAGGGGG CAAGGGAGAA GCTGCTGGTC GGAATCACA TGAAAAACGCT 60
CCTTCTTTTG CTGCTGGTGC TCCTGGAGCT GGGAGAGGCC CAAGGATCCC TTCACAGGGT 120
GCCCTCAGG AGGCATCCGT CCCTCAAGAA GAAGCTCGGG GCACGGAGCC AGCTCTCTGA 180
GTTCTGGAAA TCCCATTAAT TGGACATGAT CCAGTTCAAC GAGTCTGCT CAATGGACCA 240
15 GAGTGCCAAG GAACCCCTCA TCAACTACTT GGATATGAA TACTTCGGCA CTATCTCCAT 300
TGGCTCCCCA CCACAGAACT TCACGTGTCAT CTTCGACACT GGCTCCCTCA ACCTCTGGGT 360
CCCTCTGTG TACTGCACTA GCCCAGCCTG CAAGACGCAC AGCAGGTTC AGCCTTCCCA 420
GTCCAGCACA TACAGCCAGC CAGGTCAATC TTTCTCCATT CAGTATGGAA CCGGGAGCTT 480
GTCCGGGATC ATTGGAGCCG ACCAAGTCTC TGTGGAAGGA CTAACCGTGG TTGGCCAGCA 540
20 GTTTGGAGAA AGTGTACAG AGCCAGGCCA GACCTTTGTG GATGCAGAGT TTGATGGAAT 600
TCTGGGCGTG GGATACCCCT CCTTGGCTGT GGGAGGAGTG ACTCCAGTAT TTGACACAT 660
GATGGCTCAG AACCTGGTGG ACTTGGCGAT GTTTTCTGTC TACATGASCA GTAAACCCAGA 720
AGGTGGTGGG GGGAGGAGTG TGATTTTGG AGGCTACGAC CACTCCCAT TCTCTGGGAG 780
25 CCTGAATTG TGTCCCAAGC CCAAGCAAGC TTAATGGCAG ATTGCAGTGG ATAACATCCA 840
GGTGGGAGGC ACTGTTATGT TCTGCTCCGA GGGCTGCCAG GCCATTGTGG ACACAGGGAC 900
TTCCCTCATC ACTGGCCCTT CCGACAAGAT TAAGCAGCTG CAAAACGCCA TTGGGGCAGC 960
CCCGCTGGAT GGAGAAATAT CTGTGGAGTG TGCCAACTTT AACGTGATGC CGGATGTGAC 1020
CTTCACCATT AACGGAGTCC CCTATACCT CAGCCCAACT GCCTACACCC TACTGGACTT 1080
30 CGTGGATGGA ATGCAGTTCT GCAGCAGTGG CTTTCAAGGA CTTGACATCC ACCCTCCAGC 1140
TGGGCCCTCT TGGATCTCTG GGGATGTCTT CATTCGACAG TTTTACTCAG TTTTGTACCG 1200
TGGGAATAAC CGTGTGGGAC TGGCCCCAGC AGTCCCCATA GGAGGGGCGT TGTGTCTGTG 1260
CCTGCCGTGC TGACAGACTT TGAATATGTT AGGCTGGGGC ATTCTTTACA CCTACAAAAA 1320
GTTATTTTCC AGAATATGTA GCTGTTTCCA GGGTTGCAAC TTGAATTAAG ACCAAACAGA 1380
35 ACATGAGAA ACACACACAC ACACACATAT ACACACACAC ACATTTACA CATACACACC 1440
ACTCCACCA CGTCTATGAT GGAGGAATTA CGTTATACAT TCATATTTT TATTGATTTT 1500
TGATATGAA AATCAAAAT TTTACATTT GATTATGAAA ATCTCCAAAC ATATGCACAA 1560
GCAGAGATCA TGGTATAATA AATCCCTTGG CAATCCACT CAGCCCTGAC AACCCATCCA 1620
CACACGCCA GGCTGTGTTA TCTACACTGC TGCCCACTCC TCTCTCCAGC TCCACATGCT 1680
40 GTACCTGGAT CATCTGAGG CAAATTCGCA GCATTACATC ATTTTGTCCA TAAATATTTT 1740
TAACATCCTT AATATACAAA TCGGAATTCA AGCATCTCCC ATTGTCCCAC AAATGTTTGG 1800
CTGTTTTTGT AGTTGGATTG TTTGTATTAG GATTCAAGCA AGGCCCATAT ATTGCATTTA 1860
TTTGAATGCT CTGTAACTCT CTTTCCATCT ACAGAGTTTA GCACATTGGA ACGTTGCTGG 1920
TTGAAATCCC GAGGTGTCTT TTGACATGGT TCTCTGAACT TATCTTCCCT ATAAAAATGGT 1980
45 AGTTAGATCT GGAGGTCTGA TTTTGTGGCA AAAATACTTC CTAGGTGGTG CTGGGTACTT 2040
CTGTGTGCA CTGTGACAGA GGCAGATAAT GCTGGTGCCT CTCTATTGGT AATGTTAAGA 2100
CTGCTGGGTG GGTTTGGAGT TCTTGGCTTT AATCAITCAT TACAAAGTTC AGCATTTT

Seq ID NO: 467 Protein sequence
Protein Accession #: NP_001901.1

50 1 11 21 31 41 51
MKTLLLLLLV LLELGEAQQS LHRVPLRRHP SLKKKLRRAS QLSEFWKSHN LDMIQFTESC 60
55 SMDQSAKEPL INYLDMEYFG TISIGSPQPN FTVIPDTGSS NLWVPSVYCT SPACKTHSRF 120
QPQSSTYSQ PQQSFSIQYQ TQSLSGIIGA DQVSVEGLTV VQGQFGEVST EPGQTQVDAE 180
PDGILGLGVP SLAVGVGTPV FDNMMAQNLV DLPMPFVYMS SNEPBGAGSE LIPGSDYDHS 240
PSGLSNWVVP TKQAYWQIAL DNIQVGGTVM FCSBSCQAIV DTGTSITLGP SDKIKQLQNA 300
IGAAPVDGEY AVECANLVNM PDVTFTINGV PYTTLSPATY LLDVFDGMQP CSSGPGQLDI 360
60 HPPAGPLWLL GDVFIQFYS VPDGRNRRVG LAPAVP

Seq ID NO: 468 DNA sequence
Nucleic Acid Accession #: NM_018058.1
Coding sequence: 319..1575

65 1 11 21 31 41 51
TACCGCTGC GGGACCGGCA GGGGAACGCC ATCGGGGTCA CAGCCTGCGA CATCGACGGG 60
GACGGCCGGG AGGAGATCTA CTTCCTCAAC ACCAATAATG CCTTCTCGGG GGTGGCCACG 120
70 TACACCGACA AGTTGTTCAA GTTCCGCAAT AACCGTGGGG AAGACATCCT GAGCGATGAG 180
GTCAACGTGG CCGTGGTGTG GGCCAGCCTC TTTGCCGGAC GCTCTGTGGC CTGTGTGGAC 240
AGAAAGGGCT CTGACGCTA CTCTATCTAC ATTGCCAATT ACGCTACGG TAATGTGGGC 300
CCTGATGCC TCAATTGAAAT GGACCTGAG GCCAGTGACC TCTCCCGGG CATTCTGGCG 360
CTCAGAGATG TGGCTGCTGA GGCTGGGGTC AGCAAAATATA CAGGGGGCGG AGGCCTCAGC 420
75 GTGGGCCCCA TCCTCAGCAG CAGTGCCTCG GATATCTTCT GCGACAATGA GAATGGGCCT 480
AACTTCTTTT TCCACAACCG GGGCGATGGC ACCTTTGTGG ACGCTGCGGC CAGTGTGGT 540
GTGGACGACC CCCACGACA TGGGCGAGGT GTGCGCTGG CTGACTTCAA CCGTGTGGC 600
AAAGTGAGCA TCGTCTATGG CAACTGGAAT GGCCCCACC GCCTCTATCT GCAATGAGC 660
ACCATGGGA AGGTCCGCTT CCGGACATC GCCTCACCCA AGTCTCCAT GCCCTCCCT 720
80 GTCCGACCGG TCATCACCGC CGACTTTGAC AATGACCAAG AGCTGGAGAT CTCTTCAAC 780
AACATTGCTT ACGCGAGCTC CTCAGCCAAC CGCTCTTCC GCGTCACTCG TAGAGAGCAC 840
GGAGACCCCT TCATCGAGGA GCTCAATCCC GGCAGCGCT TGGAGCCTGA GGGCCGGGGC 900
ACAGGGGGTG TGGTGACGGA CTTGACGGA GACGGGATGC TGGACCTCAT CTGTCCCAT 960
GGAGAGTCCA TGGCTCAGCT GCTGTCCGTC TTCCGGGGCA ATCAGGGCTT CAACAACAC 1020
85 TGGCTGCGAG TGGTGCCAGC CACCCGGGTT GGGGCTTTG CCAGGGGAGC TAAGGTCTGT 1080
CTCTACACCA AGAAGAGTGG GGGCCACCTG AGGATCATCG ACGGGGGCTC AGGCTACCTG 1140
TGTGAGATGG AGCCCGTGGC ACACCTTTGGC CTGGGGAAGG ATGAAGCCAG CAGTGTGGAG 1200
GTGACGTGGC CAGATGGCAA GATGGTGAGC CGGAACGTGG CCAGCGGGGA GATGAACCTA 1260

5
 10
 15

```

GTGCTGGAGA TCCTCTACCC CCGGATGAG GACACACTTC AGGACCCAGC CCCACTGGAG 1320
ACACCAATGA ATGCATCCAG TTCCCATTCG TGTGCCCTCG AGACAAGCCC GTATGTGTCA 1380
ACACCTATGG AAGCTACAGG TGCCGGACCA ACAAGAAGTG CAGTCGGGGC TACGAGCCCA 1440
ACGAGGATGG CACAGCCTGG GTGGGGACTC TCGGCCAGTC ACCGGGCCCC CGCCCCACCA 1500
CCCCCACCGC TGCTGCTGCC ACTGCGCTG CTGCTGCCGC TGCTGGAGCT GCCACTGTCTG 1560
CACCGGTCTT CGTAGATGGA GATCTCAATC TGGGGTCGGT GGTAAAGGAG AGCTGCGAGC 1620
CCAGCTGCTG AGCAGGGGTG GGACATGAAC CAGCGGATGG AGTCCAGCAG GGGAGTGGGA 1680
AAGTGGGCTT GTGCTGCTGC CTAGACAGTA GGGATGTAAA GGCTGGGAG CTAGACCCCTC 1740
CCCAAGCCCA TCCATGCACA TTACTTAGCT AACAAATTAG GAGACTCGTA AGGCCAGGCC 1800
CTGTGCTGGG CACATAGCTG TGATCACAGC AGACAGGGTC GCTGCCCTGA TGGCGCTTAC 1860
ATTCCAGTGG GTCTAATGAC CATATCTTAG GACACAGATG TGCCAGGGA GGTGGTGTCA 1920
CTGCACAGGA AGTATGAGGA CTTTAGTGTC CTGAGTTCAA ATCCTGATTG AGGAATCAC 1980
AAAGCTATGT TACCTTACAG CAGTCACTTA ACTTGTAGC CATCCATTAT CGCATCTGCA 2040
AAATGGGARG ARVLYTKKS GAHLRIIDGG SGYLCMEFV AHFGLKDEA SSVEVTWPDG 2100
GACACTTGGC ACAAACCTG GCACATAGTA AAGGCTCAAT AAAACAAGT GCCTCTCACT 2160
GGGCTTTGTC AACACGTG
  
```

Seq ID NO: 469 Protein sequence
 Protein Accession #: NP_060528.1

20
 25
 30

```

1 11 21 31 41 51
MDPEASDLR GILALRDVAA EAGVSKYTG RGVSVPILS SSASDIFCDN ENGNPFLFHN 60
RGDGTFDVAA ASAGVDDPHG HGRGVALADF NRDKVDIVY GNMNGPHRLY LQSTHGKVR 120
FRDIASPKFS MPSPVRTVIT ADPDNDQELE IFFNNIAYRS SSANRLFRVI RREHGDPLIE 180
ELNPGDALEP EGRGTGGVVT DFDGDMGLDL ILSHGSEMAQ PLSVFRGNQG FNNNMLRVVP 240
RTRVGAFARG ARVLYTKKS GAHLRIIDGG SGYLCMEFV AHFGLKDEA SSVEVTWPDG 300
KMSRNVASG ENNSVLEILY PRDEDTLQDP APLETPMNAS SSHSCALET S PYVSTFMEAT 360
GAGPTRSAVG ATSPTRMAQP AWGLSASHRA PAPPPPPPLL PLPLLLPLLE LPLLHRSS
  
```

Seq ID NO: 470 DNA sequence
 Nucleic Acid Accession #: AJ279016
 Coding sequence: 1..1962

35
 40
 45
 50
 55
 60
 65
 70
 75

```

1 11 21 31 41 51
ATGTCCAGGA TGTACCGTT CTGTGCTGCTG CTCTGGTTTC TGCCCATCAC TGAGGGGTCC 60
CAGCGGCTG AACCCTAGTT CACTGCAGTC ACCAACTCAG TTCTGCCTCC TGACTATGAC 120
AGTAATCCCA CCCAGCTCAA CTATGGTGTG GCAGTTACTG ATGTGGACCA TGATGGGGAC 180
TTTGAGATGG TGTGGGGGG GTACAATGGA CCCAACCTGG TTCTGAAGTA TGACCGGGCC 240
CAGAAGCGGC TGGTGAACAT CGCGGTGATG GAGCGCAGCT CACCCTACTA CGCGCTGCGG 300
GACCGGCGAG GGAACGCCAT CGGGGTGACA GCCTGCGACA TCGACGGGGA CGGCGGGGAG 360
GAGATCTACT TCTTCAACAC CAATAATGCC TTCTCGGGGG TGCCACGTA CACCGACAAG 420
TGTTCRAAGT TCCGCAATAA CCGGTGGGAA GACATCCTGA GCGATGAGGT CAACGTGGCC 480
CGTGGTGTGG CAGCGCTTCT TGCCGAGCGC TCTGTGGCCT GTGTGGACAG AAAGGGCTCT 540
GGAGCTACTT CTATCTACAT TGCCAATTAC GCCTACGGTA ATGTGGGGCC TGATGCCCTC 600
ATTGAATGG ACCCTGAGGC CAGTGACCTC TCCCGGGGCA TTCTGGCGCT CAGAGATGTG 660
GCTGCTGAGG CTGGGGTCAG CAATATACA GGGGGCGAG GCGTCAGCGT GGGCCCCATC 720
CTCAGCAGCA GTGCTCGGGA TATCTTCTGC GACAATGAGA ATGGGCCTAA CTTCCTTTTC 780
CAACAACGGG GCGATGGCAC CTTTGTGGAC GCTGCGGCCA GTGCTGGTGT GGACGACCCC 840
CACCAGCATG GCGGAGGTGT CGCCCTGGCT GACTTCAACC GTGATGGCAA AGTGGACATC 900
GTCTATGGCA ACTGGAATGG CCCCCACG CTTATCTGCT AAATGAGCAC CCATGGGAAG 960
GTCCGCTTCC GGGACATCGC CTCACCCAAG TTCTCCATGC CCTCCCCTGT CCGCACGGTC 1020
ATCACCGCGG ACTTTGACAA TGACCAGGAG CTGGAGATCT TCTTCAACAA CATTGCCTAC 1080
CGCAGCTCCT CAGCCAACCG CCTCTTCGCT GTCATCCGTA GAGAGCACGG AGACCCCTC 1140
ATCGAGGAGC TCAATCCCGG CGAOGCCTTG GAGCCTGAGG GCGGGGGCAC AGGGGGGTGT 1200
GTGACCGACT TCGACGGAGA CCGGATGCTG GACCTCATCT TGTCCCATGG AGAGTCCATG 1260
GCTCAGCGCT TGTCCGCTT CCGGGGCAAT CAGGGCTTCA ACAACAAGT GCTGCGAGTG 1320
GTGCCACGCA CCGGTTTGGG GGCCTTTGCC AGGGGAGCTA AGGTGCTGCT CTACACCAAG 1380
AAGAGTGGGG CCGCCTGAGG GATCATCGAC GGGGGCTCAG GCTACCTGTG TGAGATGGAG 1440
CCCGTGGCAC ACTTTGGCCT GGGGAAGGAT GAAGCCAGCA GTGTGGAGGT GACGTGGCCA 1500
GATGGCAAGA TGGTGAGCCG GAACGTGGCC AGCGGGGAGA TGAATCAGT GCTGGAGATC 1560
CTCTACCCCC GGGATGAGGA CACACTTCAG GACCCAGCCC CACTGGAGTG TGGCCAAGGA 1620
TTCTCCAGC AGGAAATGG CCAATTGCATG GACACCAATG AATGCATCCA GTTCCCATTG 1680
GTGTGCCCTC GAGACAAGCC CGTATGTGTC AACACCTATG GAAGCTACAG GTGCCGAGAC 1740
AACAGAAGT GCAGTCGGGG CTAAGAGCCC AACGAGGATG GCACAGCCTG CGTGGGGACT 1800
CTCGGCCAGT CACCGGGCCC CGGCCCCACC ACCCCACCGC CTGCTGCTGC CACTGCCGCT 1860
GCTGCTGCGG CTGCTGGAGC TGCCACTGCT GCACCGGTCC TCGTAGATGG AGATCTCAAT 1920
CTGGGGTCTG TGGTTAAGGA GAGCTGCGAG CCCAGCTGCT GAGCAGGGGT GGGACATGAA 1980
CCAGCGGATG GAGTCCAGCA GGGGAGTGGG AAAGTGGGCT TGTGCTGCTG CCTAGACATG 2040
AGGATGTAA AGGCTCGGGA GCTAGACCTT CCCCAGCCCC ATCATGCAC ATTACTTAGC 2100
TAACAATTAG GGAGACTCGT AAGGCCAGGC CCGTGTGCTG GCACATAGCT GTGATCACAG 2160
CAGACAGGGT CGCTGCCCTT ATGGCGCTTA CATTCCAGTG GGTCTAATGA CCAATCTTCA 2220
GGACACAGAT GTGCCCGGG AGGTGGTGTG ACTGCACAGG AAGTATGAGG ACTTTAGTGT 2280
CCTGAGTTCA AATCCTGATT CAGGAACCTA CAAAGCTATG TGACCTTACA CCACTCACTT 2340
AACTGTGTAG CCATCCATTA TCGCATCTGC AAAATGGGGA TTAAGAAATG AATCTTGGGG 2400
TTAGTGTGGA GATTAGATT AATGTATGTA AGACACTTGG CACAAAACCT GGCACATAGT 2460
AAAGGCTCAA TAAAAACAAG TGCGCTTCTC CAACAG
  
```

Seq ID NO: 471 Protein sequence
 Protein Accession #: CAC08451

80
 85

```

1 11 21 31 41 51
MSRMLPFLLL LNFLPITEGS QRAEPMFTAV TNSVLPFDYD SNPTQLNYGV AVTDVDHGDG 60
FEIVVAGYNG FNLVLKYDRA QKRLVNIADV ERSSPYALR DRQNAIGVT ACDIDGDGRE 120
EIVFLTNNA FSGVATYTDK LFKFRNRWE DILSDEVNVA RGVASLFAGR SVACVDRKGS 180
  
```

GRYSIIYANY AYGNVGPDAL IEMDPEASDL SRGILALRDV AAEAGVSKYT GGRGVSVGPI 240
 LSSASDIFC DNENGNPFLF HNRGDGTFVD AAASAGVDDP HQHGRGVALA DFNRDGKVDI 300
 VYGNWNGPHR LYLQMSHKGK VRFRDIASPK FSPSPSPVTV ITADFDNDQE LEIPFNNIAY 360
 RSSSANRLFR VIRREHGDPL IEELNPGDAL EPEGRGTGGV VTDPDGDMGL DILSHGESHM 420
 AQLPSVFRGN QGFNNWNLVR VPRTRFGAFA RGAKVVLVYTK KSGAHLRIID GGSVYLCEME 480
 PVAHPLGLKD BASSVEVWTF DGKMSVSRVA SGEMNSVLEI LYPRDEDTLQ DPAPLECOQG 540
 PSQENGHGM DTNECIQFPF VCPDRKPVCV NTYGSYRCRT NKKCSRGYEP NEDGTACVGT 600
 LGQSPGPRFT TPTAAATAAA AAAAAGAATA APVLVDGDLN LGSVVKESCE PSC

10 Seq ID NO: 472 DNA sequence
 Nucleic Acid Accession #: FGENESH
 Coding sequence: 1..4794

15 1 11 21 31 41 51
 | | | | | |
 ATGGCGTGTG CGGAGGAGCT CCCAGCCCGT TGCTCTGGTT GGATGGGACT GGGTGGGCCC 60
 AGCGGCTCCT CCCAGCATC CCCTCCCAT TCCTCTCCA GGTACAATGG ACCCAACCTG 120
 GTTCTGAAGT ATGACCGGGC CCAGAAGCGG CTGGTGAACA TCGCGTCTGA TGAGCGCAGC 180
 TCACCCCTACT ACGCGCTGCG GGACCGGCAG GGGAAACGCC TCGGGGTAC AGCCTGCGAC 240
 20 ATCGACGGGG ACGGCCGGGA GGAGATCTAC TTCTCAACA CCAATAATGC CTCTCGGGC 300
 CACAGCAGCT CAGCGCAGGT CCCTCTGGG CTCCACAGAA ACAGGCTGT GCTGAAGCCT 360
 CCACCTACAA CCCCTGCAGG CCTCTGGGT CTGCTCCAC TCAGCGGAAG GGACTTTTCC 420
 TCCTCCCTGG GTCAAGCTTC TCCGGACAGC AGGCAGGGAG AGAGGGTGC GGTTCCTGTC 480
 25 TGTCGGGGT GACTGAGACC TACCCATGAA CCGAACCAT TTCTCTGAG ACCCAATCA 540
 GGGGTGGCCA CGTACACCA CAAGTTGTTT AAGTTCGCA ATAACCGTG GGAAGACATC 600
 CTGAGCGATG AGGTCAACGT GGCCCGTGGT GTGGCCAGCC TCTTTGCCG ACCTCTGTC 660
 GCCTGTGTGG ACAGAAAGGG CTCTGGACGC TACTCTATCT ACATTGCCAA TTACGCTTAC 720
 GGTAATGTGG GCCCTGATGC CCTCATGAA ATGACCCCTG AGGCCAGTGA CCTCTCCCG 780
 30 GGCAATCTGG CGCTCAGAGA TGTGGCTGCT GAGGCTGGGG TCAGCAATA TACAGAAGGC 840
 TTCTCCACA CTGCTCTCC AAGCATTGGT GAGATATCTG GCAGAACCGA GGAGCGGGAA 900
 GGAGGAGACG CAGAGGAGGC AGATGAGGAG CACAGTGGGG ATGGAAGCAC CAGCCAACTG 960
 TGCCCGCTGG GCTGGAGAGA CGGGCAGTTC AAGGAAGAAG CAGCAGCTTT GGTGGAGGAA 1020
 CAGAGGGAGG CTGGGCGAGC TGGCGTGCCC AGAGGACGTG TTGGAACAGC TCTGCAGACT 1080
 35 TCCAAAGAGC ATTTGGCTGA CAAGAACCTA TTTGGCCAC CATGTTACTA TTCTGTCTGC 1140
 GCGCCTTCTC CAGCCACGCC TTTCCCTGCC CGCCAAGCCC CCCAACCTA CCTGTAGACC 1200
 CCCTTGTGCA CTGAGCTAAT GACACATGGA CGTCTGGCTG GAAAACTAGC CCGGAGTGTG 1260
 CCCACCCGCC GAGCCCGCAG AATGGACCCC AAATGTAAAG GCGCCATGC TGAGCCCGGC 1320
 CTGATGGCTG AGGCTTTGGG CGCGTGCCCA GCGCTCAGCA CCACTGTGGT GCCAGGGGGC 1380
 40 CTGAGAAGCT GGGAGGAAAG CAGGCAGAA GGGCAGGCCA TGTCAGATG TGCACTCAGG 1440
 GAGCTGGGAG GTCCCTGGAG CCAAGCCACA CAGCACCTGC CTGCTAGAGA GCTGTATGAC 1500
 CTGGGAGAAC CTCCATTTT ACAAGAACCA GACGAGATC CAGGGAGGAG AAGGGACTCG 1560
 CCCAAGTCA CACAGAGATG CCATCTAGT GCCACCATGC CAGCTCTCGG GGGACTCGAG 1620
 GGCCCCGGGA GGTGGCCAA GCGAGAGATT GGGAGAGAGA CTGGGCGAGT AGGAAGACCA 1680
 45 CTCTCCCATC CCTTGGTCCC CAACTTCCCC AGCTGCTTGA GGCCTCTTGA AGCCCGGACA 1740
 GTGCGGGGAG CTGCGCTGCC TGGGAATCCT GGGAACTGGG TTCTGGACAT GGCCAAGGCC 1800
 CTGGCTGGA ACCAGATGGA AAAAGAGGAG GGGAAAGTTC ATGGAAGACA TGAGCCAGCA 1860
 TTTAGGCTCA GGAAGCAGG GGAAGCAGAA TTCCCCCAG GCTCCTCTGA GGAGCCTCTG 1920
 CTGCAATTCC CCTCAGGCTC CAGAGGCAGC CCTGTCTTCC AGGTGGGCTT GGGCTTGTCT 1980
 50 TCTGCCACTC ACTGTGGGTC GATGTCTTTT CTAGGGGGCC GAGGCGTCAG CGTGGGCCCC 2040
 ATCCTCAGCA CAGTGCCTTC GGATATCTTC TGCGACAATG AGAATGGGCC TAACTCTCTT 2100
 TTCCACAACC GGGCGGATGC CACCTTTGTG GACGCTCGCG CCACTGTCTGA ACCTGCTTTA 2160
 GCCTTCTATG TTCACTCAA ATATCACCTC TGCAGAGATT TTCTCACTC CCGTGTGCCAC 2220
 CTAGCAGAAA CTGCTCTTTC CTCTCTCTGC TGCCCGTGGC ATGCACTGCT TCTTCAGGCT 2280
 55 CCACATTGCC ATCATGGTTT GTCTATGAGC TTTACAAGGA CCGGGTCAGG GTTCTATTCA 2340
 TTCTGAAGC AAGGCTTGGC CTCCAGTGCC CACCGGAGGA CACTCAGCCT CCAGGGTTCT 2400
 CAGGGGGGCC CACCTGCTCT TCTGGCAAGA GCTCCCTGTG TCTGGGGTCT TCTGATCCCC 2460
 ACTGCTTATT ACAATTGCTT GTGGTCTGCC ATCCAGAGA GCTGTATGAC CCACAGCTAT 2520
 TTGTCTCTG AAAGAGTCAA CGTGGGTGTG GACGACCCCC ACCAGCATG GCGAGGTGTC 2580
 60 GCCTGGGTG ACTTCAACCG TGATGGCAAA GTGACATCG TCTATGGCAA CTGGAATGGC 2640
 CCCCACGGCC TCTATCTGCA AATGAGCACC CATGGGAAGG TCCGCTTCCG GGCATCTGCC 2700
 TCACCCAGT TCTCCATGCT CTCCCTGTG CCGACGGTCA TCACCGCGA CTTTGACAA 2760
 GACCAGGAG TGGAGATCTT CTTCAACAAC ATTGCTTACC GCAGCTCTCT AGCCAAACGC 2820
 CTCTCCGAT GCTCCATCTT GGCTCGTGGC TCTTCATCTT TGACAGCTG TGGAGGAAC 2880
 65 GGTGAGGAG AAGGTTTAAG AATCAGAAAG GAGGGGTTC CAGGGCCAGG GGGTCAGGCC 2940
 AAGGTCAACA CAGTCCCTCT GATGAAGAAA CAGAAAGGAA GGAAGGACGA GGAAGTGGCA 3000
 AGAGGCTGTG GGAATGCAGG GCAAAGCCTG GCCAAGGAGC CGGCTCTGCT TATTGAGGG 3060
 AAAGGGAAGG GAATGTGGC CCAAAGTGTG CCCAGAACCC AAGCGCACCA AGATACAAAG 3120
 CCACACTACC ACAAAAGGG GCTACAGGT CCAATCACTA CCAGGAAAG GGGCTACGGG 3180
 70 GTCCAATCAC TACCAGGAAA AGGGGCTACG GGGTCCAATC ACTACAGGA AAAGGGGCTA 3240
 CGGGGTCCAA TCACTACCA GAAAAGGGGC TACGGGGTCC AATCACTACC AGGAAAAGGG 3300
 GCTACGGGCT CCAATCACTA CCAGGAAAAG GGGCTACAGG GTCCAATCAC TACCAGGAAA 3360
 AGGGGCTACG GGTCCAATC ACTACAGGA AAAGGGGCTA CAGGGTCCAA TCACTACCA 3420
 75 AGAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGGG GCTACGGGT CCAATCACTA 3480
 CAGGAAAGG GGGCTACAGG GTCCAATCAC TACCAGGAAA AGGGGCTACG GGTCCAATC 3540
 ACTACAGGA AAAGGGGCTA CGGGCTCCAA TCACTACCA GAAAAGGGGC TACGGGGTCC 3600
 AATCACTACC AGGAAAAGGG GCTACAGGT CCAATCACTA CCAGGAAAAG GGGCTACAGG 3660
 GTCCAATCAC TACCAGGAAA AGGGGCTACG GGTCCAATC ACTACAGGA AAAGGGGCTA 3720
 80 CGGGGTCCAA TCACTACCA GAAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGGG 3780
 GCTATGGGT CCAATCACTA CCAGGAAAAG GGGCTACAGG GTCCAATCAC TACCAGGAAA 3840
 AGGGGCTATG GGTTCCAATC ACTACCAAG AAAGGGGCTA CGGGGTCCAA CGTCACTCGT 3900
 AGAGAGCAG GAGACCCCT CATCGAGGAG CTCATCCCG GCGACGCTT GAGCCTGAG 3960
 GGCCTGGGCA CAGGGGTGT GGTGACCGAC TTCAGCGGAG ACGGATGCT GACCTCATC 4020
 85 TTGTCCATG TGTCCGTC GTGTCAGCG CTGTCCTGT TCGGGGCAA TCAGGGCTTC 4080
 AACCAACTG GGTGCGAGT GGTGCCACGC ACCCGGTTG GGGCTTTGC CAGGGGAGCT 4140
 AAGGTGCTG TCTACCCAA GAAGAGTGGG GCCCACTGA GGATCATCGA CGGGGGCTCA 4200
 GGTACTCTGT GTGAGATGGA GCCCGTGGCA CACTTTGGCC TGGGGAAGGA TGAAGCCAGC 4260
 AGTGTGGAG TGAGCTGGCC AGATGGCAAG ATGTGAGGCC GGAACGTGGC CAGCGGGGAG 4320

WO 02/086443

PCT/US02/12476

5

ATGAACCTCAG TGCTGGAGAT CCTCTACCCC CGGGATGAGG ACACACTTCA GGACCCAGCC 4380
CCACTGGAGT GTGGCCAAAG ATTCTCCAG CAGGAAAATG GCCATTGCAT GGACACCAAT 4440
GAATGCATCC AGTTCCCAT CTGTGCCCCT CGAGACAAGC COGTATGTGT CAACACCTAT 4500
GGAAAGCTACA GGTGCGGAC CAACAAGAA GTCATCGGG GCTACGAGCC CAACGAGGAT 4560
GCCACAGCCT GCGTGGGTAC TGAGCTAGGC TCTAGGCATA CAATGACGTG GAAACCAAGG 4620
CCCAAAAAGG AGCTGCAACT TTCCAAGGC ATCTGCACCC CCGTCTGGTC CTTTTCTCTG 4680
CCGGGTGCC GCGTCTCCT CAAAAGAGCT CAGCTCCAGG CTGCTCCAG CACCTTCTC 4740
CAGAAAGCTC CAGGTATTCC AGAAGCCCAA GTGTATGAAC AAGATCAGGA ATAA

10

Seq ID NO: 473 Protein sequence
Protein Accession #: FGENESH predicted

15

1 11 21 31 41 51
MACPGLPAR CSGMGLGPP SSSSPASPPH SSSRYNGPFL VLKYDRAQKR LVNIAVDERS 60
SPFYALDRQ GNAIGVTACD IDGDREEIY FLNTNNAFSG HSSSAQVPSG LHRNRPVLKP 120
PPTTPAGLIG LPPLSGRDFS SSLQASPD S RQGERVPVPC CRGGLRP THE PEPFLLRPKS 180
GVATYTDKLF KFRNRWEDI LSDEVNVARG VASLFAGRSV ACVDRKSGSR YSIYIANYAY 240
GNVGPALIE MDPEASDLR GILALRDVAA EAGVSKYTEG FSHTASPSIG EISGRTEERE 300
GGDPEEADDE HSGDGSSTQL CRLGWKDGQF KEEAAALVEE QREAGAAGVP RGRVRTALQT 360
SKSHLADKLN FGPPCYISVC APSPAHPPA RQAPQHYPVA PLVTQLMTHG RLAGKLARSV 420
PHPRAPGMD KCKGRHAEPG LMAEALGAWP ALSTTVVPGG LRSWEESRQK GQAMSRCALR 480
ELGGPWSQAT QHLPALELYD LGPEPILQRT DGDGRRRDS PKVTQECHLV ATMPALGGLE 540
GPGRVAKREI GRETPAVGRP LSHPLVPNFP SCLRPLEAGT VPGAALPGNP GNWVLDMAKA 600
LAWNQMEKEE GKIHGDHEPR FRLRKAREAE FPPGSSEEP LQFPSSGLRGS PVLQVGLGLA 660
SATHCGSMSF LGGRGVSVGP ILSSASDIF CDNENGPNFL FHNRGDGFV DAAASAERRL 720
AFIVHLKYLH CRDPFPHSLCH LAETGPSSSC CPWHARLLQA PHCHGLSMS FRTGSRFYS 780
FLTQGLASSA HRTLSLQGS QGAPPCLLAR APCVLGSLIP TAYYIVLWSA IPESLMTHSY 840
LSSERVNVGV DDPHQHGVRP ALADFNRDGK VDIVYGNWNG PHRLYLQMSI HGKVRFRDIA 900
SPKFSMPSPV RTVITADFDN DQELEIFNN IAYRSSANR LFRCSILARG SSSLTAGGRN 960
QGEGRLRIR GGFPGPGQQA KWTGPGPLMK QKRGRDEDA RCGNAGQSL AKEPASAIA 1020
KGKGNVAQSV PRTQAPQDTK PHYHKKGLQG PITTRRGYV VQSLPGKGT GSNHYQKGL 1080
RGPITTRKRG YGVQSLPGK ATGSNHYQEK GLQGPITTRK RGYGLQSLPG KGATGSNHYH 1140
RKGLRAPITT RKRGYGVQSL PGKATGSNH YQEKGLRGI TTRKRGVGLP SLPGKATGS 1200
NHYQEKGLQG PITTRRGYV VQSLPGKAT GSNHYQEKGL RGPITTRKRG YGLQSLPGKE 1260
AMGSNHYQEK GLRAPITTRK RGYGVQSLPG KGATGSNVIR REHGDPLIEE LNPGDALPE 1320
GRGTGGVVD FDGDMLDLI LSHGESMAQP LSVFRGNQGF NNNWLRVVRP TRFGAFARGA 1380
KVVLVTKKSG AHLRIIDGGS GYLCEMEFVA HFGLGKDEAS SVEVTWPDGK MVSRRVAVSGE 1440
MNSVLEILYP RDEDTLQDPA PLECGQGFSG QENGHCMDTN ECIQFPFVCP RDKPVCVNTY 1500
GSYRCRTNKK CSRGYEPNED GTACVGTGLG SRHTWTWKPR PKKELQLSQG ICTPVWSFFL 1560
PGCRLLLKRA QLQAPSTLL QKAPGIPBAQ VYEQDQE

45

Seq ID NO: 474 DNA sequence
Nucleic Acid Accession #: NM_003661.1
Coding sequence: 1..1152

50

1 11 21 31 41 51
ATGAGTGCAC TTTTCTTGG TGTGGGAGTG AGGGCAGAGG AAGCTGGAGC GAGGGTGCAA 60
CAAAACGTTT CAAGTGGGAC AGATACTGGA GATCCTCAAA GTAAGCCCTT CGGTGACTGG 120
GCTGCTGGCA CCATGGACCC AGAGAGCAGT ATCTTTATTG AGGATGCCAT TAAATATTTC 180
AAGAGAAAAG TGAGCACACA GAATCTGCTA CTCCTGCTGA CTGATAATGA GGCCTGGAAC 240
GGATTGCTGG CTGCTGCTGA ACTGCCAGG AATGAGGCAG ATGAGCTCG TAAAGCTCTG 300
GACAACTTGT CAAGACAAAT GATCATGAAA GACAAAACT GGCACGATAA AGGCCAGCAG 360
TACAGAAACT GTTTCTGAAT AGAGTTTCCT CGGTGAAAA GTGAGCTTGA GGATAACATA 420
AGAAGGCTCC GTGCCCTTGC AGATGGGGTT CAGAAGGTCC ACAAGGCAC CACCATCGCC 480
AATGTGGTGT CTGGCTCTCT CAGCAITTC TCTGGCATCC TGACCTCGT CGGCATGGGT 540
CTGGCACCTT TCACAGAGGG AGGCAGCCTT GTACTCTTGG AACCTGGGAT GGAGTTGGGA 600
ATCACAGCCG CTTTGACCGG GATTACCAGC AGTACCATGG ACTACGAAA GAAGTGTGTG 660
ACACAAGCCC AAGCCCAACGA CCTGGTCATC AAAAGCCTTG ACAAAATTGA GGAGGTGAGG 720
GAGTTTTTGG GTGAGAACAT ATCCAACTTT CTTTCTTAG CTGGCAATAC TTACCAACTC 780
ACACGAGGCA TTGGGAAGGA CATCCGTGCC CTCAGACGAG CCAGAGCCAA TCTTCAGTCA 840
GTACCGCATG CCTCAGCCTC ACGCCCCCGG GTCACTGAGC CAATCTCAGC TGAAGCGGT 900
GAACAGGTGG AGAGGGTTAA TGAACCCAGC ATCTCGAAA TGAGCAGAGG AGTCAAGCTC 960
ACGGATGTGG CCCCTGTAAG CTCTTCTTCT GTGCTGGATG TAGTCTACCT CGTGTACGAA 1020
TCAAGCACT TACATGAGGG GGCAGAGTCA GAGACAGCTG AGGAGCTGAA GAAGGTGGCT 1080
CAGGAGCTGG AGGAGAAGCT AAACATTCTC AACAATAATT ATAAGATTCT GCAGGCGGAC 1140
CAAGAAGTGT GA

70

Seq ID NO: 475 Protein sequence
Protein Accession #: NP_003652.1

75

1 11 21 31 41 51
MSALFLGVGV RAEEAGARVQ QNVPSGTDGT DPQSKPLGDW AAGTMDPES IFIEDAIKYF 60
KEKVSTQNL LLLTDNEAWN GFVAAELPR NEADELRKAL DNLARQIMK DKNWHDKQQ 120
YRNWFLKEFP RLKSELEDNI RRLRALADGV QVHKGTTIA NVVSGSLIS SGILTLVGMG 180
LAPFTEGSSL VLLEPGMELG ITAALTGITS STMDYGKKNW TQAQADLVI KSLDKLKEVR 240
EFLGENISNF LSLAGNTYQL TRGIGKDIRA LRRARANLQS VPHASASRPR VTEPISAESG 300
EQVERVNEPS ILEMRSVKVL TDVAPVSFFL VLDVVYLVE SKHLHEGAKS ETAELKKVA 360
QELEEKLNL NNNYKLLQAD QEL

85

Seq ID NO: 476 DNA sequence
Nucleic Acid Accession #: NM_014452.1
Coding sequence: 1..1968

1 11 21 31 41 51

	ATGGGGAGCT	CTCCGAGCAG	CAGCACCGCC	CTCGCTCTCT	GCAGCCGCAT	CGCCCCCGGA	60
	GCCACAGCCA	CGATGATGCG	GGGCTCCCTT	CTCCTGCTTG	GATTTCCTTAG	CACCACCACA	120
5	GCTCAGCCAG	AACGAAGGG	CTCGAATCTC	ATTTGGACAT	ACCGCCATGT	TGACCGTGCC	180
	ACCGGCGAGG	TGCTAACCTG	TGACAAGTGT	CCAGCAGGAA	CCTATGCTCT	TGAGCATGTG	240
	ACCAACACAA	GCTTGCGCGT	CTGCAGCAGT	TGCCCTGTGG	GGACCTTTAC	CAGGCATGAG	300
	AATGGCATAG	AGAAATGCCA	TGACTGTAGT	CAGCCATGCC	CATGGCCAAT	GATTGAGAAA	360
	TTACCTTTGT	CTGCCTTGAC	TGACCGAGAA	TGCACCTTGC	CACCTGGCAT	GTTCCAGTCT	420
10	AACGCTACCT	GTGCCCCCCA	TACGGTGTGT	CCTGTGGGTT	GGGGTGTGCG	GAAGAAAGGG	480
	ACAGAGACTG	AGGATGTGCG	GTGTAAGCAG	TGTGCTGGGG	GTACCTTCTC	AGATGTGCGT	540
	TCTAGTGTGA	TGAAATGCCA	AGCATACACA	GACTGTCTGA	GTCAGAACCT	GGTGGTGATC	600
	AAGCCGGGGA	CCAAGGAGAC	AGACAACGTC	TGTGGCACAC	TCCCGTCTCT	CTCCAGCTCC	660
	ACCTCACTTT	CCCCCTGGAC	AGCCATCTTT	CCACGCCCTG	AGCAGATGGA	AACCCATGAA	720
15	GTCCCTTCTT	CCACTTATGT	TCCCCAAGGC	ATGAACTCAA	CAGAATCCAA	CTCTTCTGCC	780
	TCTGTAGAG	CAAAGGTACT	GAGTAGCATC	CAGGAAGGGA	CAGTCCCTGA	CAACACAAGC	840
	TCAGCAAGGG	GGAGGAAGA	CGTGAACAAG	ACCTTCCCAA	ACCTTCAGGT	AGTCAACCCAC	900
	CAGCAAGGCC	CCCAAGGAGC	ACACATCTTG	AAGCTGCTGC	CGTCCATGGA	GGCCACTGGG	960
	GGCGAGAAAT	CCAGCAGCCG	CATCAAGGGC	CCCAAGAGGG	GACATCTTAG	ACAGAACCTA	1020
20	CACAAGCAAT	TGTGACATCA	TGAGCATTTG	CCCTGGATGA	TGTGCTTTT	CCTGCTGCTG	1080
	GTGCTTGTGG	TGATTGTGGT	GTGCAGTATC	CGGAAAAGCT	CGAGGACTCT	GAAAAGAGGG	1140
	CCCCGGCAGG	ATCCAGATGC	CATTGTGGAA	AAGGAGGGC	TGAAGAAATC	CATGACTCCA	1200
	ACCCAGAAAC	GGGAGAAATG	GATCTACTAC	TGCAATGGCC	ATGGTATCGA	TATCCTGAAG	1260
	CTTGTAGACG	CCCAAGTGGG	AAGCCAGTGG	AAAGATATCT	ATCAGTTTCT	TTGCAATGCC	1320
25	AGTGAGAGGG	AGGTTGCTGC	TTTCTCCAAT	GGGTACACAG	CGGACCACGA	CGGGGCTTAC	1380
	GCAGCTCTGC	AGCACTGGAC	CATCCGGGGC	CCCGAGGCCA	GCCTCGCCCA	GCTAATTAGC	1440
	GCCCTGCGCC	AGCACCGGAG	AAACGATGTT	GTGGAGAAGA	TTGTTGGGCT	GATGGAAGAC	1500
	ACCACCCAGC	GGGAAACTGA	CAAACTAGCT	CTCCCGATGA	GCCCCAGCCC	GCTTAGCCCG	1560
	AGCCCCATCC	CCAGCCCCAA	CGCGAACTT	GAGAATTCOG	CTCTCCTGAC	GGTGGAGCCT	1620
30	TCCCCCAGG	ACAGAACAAC	GGGCTTCTC	GTGGATGAGT	CGGAGCCCTT	TCTCCGCTGT	1680
	GACTCTACAT	CCAGCGGCTC	CTCCGCGCTG	AGCAGGAACG	GTTCTTTTAT	TACCAAGAA	1740
	AAGAAGGACA	CAGTGTGTGG	GCAGGTAGCG	CTGGACCCCT	GTGACTTGCA	GCCTATCTTT	1800
	GATGACATGC	TCCACTTTCT	AAATCCTGAG	GAGCTGGGGG	TGATTGAAGA	GATTCCCGAG	1860
	GCTGAGGACA	AACTAGACCG	GCTATTGCAA	ATTATTGGAG	TCAAGAGCCA	GGAAGCCAGC	1920
35	CAGACCTCC	TGGACTCTGT	TTATAGCCAT	CTTCTGACC	TGCTGTAG		

Seq ID NO: 477 Protein sequence

Protein Accession #: NP_055267.1

	1	11	21	31	41	51	
40	MGTSPPSSSTA	LASCSRIARR	ATATMIAGSL	LLLGLFLSTT	AQPEQKASNL	IGTYRHVDRA	60
	TGQVLTCDC	KPAGTYVSEHC	TNTSLRVCS	CPVGTFTRE	NGIEKCHDCS	QPCPWFMIK	120
	LPCAALGDRE	CTCPFGMFQS	NATCAPHTVC	PVWGVRKKG	TETEDVRCKQ	CARGTFSDVP	180
	SSVMKKXAYT	DCLSQNLVVI	KPGTKETDNV	CGTLPSFSSS	TSPPSGTAIF	PREHEMETHE	240
45	VPSSTYVPKG	MNSTESNSSA	SVRPKVLSSI	QEGTVPDNTS	SARGKEDVNK	TLPNLQVVNH	300
	QQGPHHRHIL	KLLPSMEATG	GEKSSTPIKG	PKRGHPRQNL	HKHFDINEHL	PWIMVLFLL	360
	VLVVIVVCSI	RKSRTLLKKS	PRQDPSAIVE	KAGLKKSMTP	TQNRKWIYY	CNGHGIDILK	420
	LVAAGVGSQW	KDIYQFLQNA	SEREVAAFSN	GYTADHERAY	AALQHWITIR	PEASLAQLIS	480
	ALRQHRND	VEKIRGLMED	TTQLETDKLA	LPMSPSPLSP	SPIPSFNAKL	ENSALLTVEP	540
50	SPQDKNKGFF	DESEPLLR	DSTSSGSSAL	SRNGSFITKE	KKDTVLRLQR	LDPCDLQPIF	600
	DDMLEPLNPE	ELRVIEIIPQ	AEDKLDRLFE	IIGVKSQEAS	QTLDSVYSH	LPDLL	

Seq ID NO: 478 DNA sequence

Nucleic Acid Accession #: XM_044533

Coding sequence: 238..2751

	1	11	21	31	41	51	
60	GCTCTGCCCA	AGCCGAGGCT	GCGGGGCGGG	CGCCGGCGGG	AGGACTGCGG	TGCCCCGCGG	60
	AGGGGCTGAG	TTTGCCAGGG	CCCACTTGAC	CCTGTTTCCC	ACCTCCCGCC	CCCCAGGTCC	120
	GGAGGCGGGG	GCCCCCGGGG	CGACTCGGGG	GCGGACCGCG	GGGCGGAGCT	GCGCCCGGTG	180
	AGTCGGCGCG	AGCCACCTGA	GCCCGAGCCG	CGGACACCG	TGCTCTCTGC	TCTCGAATG	240
	CTGCGCACCG	CGATGGGCGT	GAGGAGCTGG	CTCGCCGCCC	CATGGGGCGC	GCTGCCGCTT	300
	CGGCCACCGC	TGCTGCTGCT	CCTGCTGCTG	CTGCTCTGCT	TGCAGCCGCC	GCTCCGACC	360
65	TGGGCGCTCA	GCCCCCGGAT	CAGCCTGCCT	CTGGGCTCTG	AAGAGCGGCC	ATTCTCTAGA	420
	TTGGAAGCTG	AACACATCTC	CAACTACACA	GCCCCCTTGC	TGAGCAGGGA	TGGCAGGACC	480
	CTGTAGCTGG	GTGCTCGAGA	GGCCCTCTTT	GCACTCAGTA	GCAACCTCAG	CTTCTGCGCA	540
	GGCGGGGAGT	ACCAGGAGCT	GCTTTGGGGT	GCAGACGCG	AGAAGAAACA	GCAGTGCAGC	600
	TTCAAGGGCA	AGGACCCACA	GCGCGACTGT	CAAAACTACA	TCAAGATCCT	CCTGCCGCTC	660
70	AGCGGCAGTC	ACCTGTTTCA	CTGTGGCACA	GCAGCCTTCA	GCCCCATGTG	TACCTACATC	720
	AACATGGAGA	ACTTCAACCT	GGCAAGGGAC	GAGAAGGGGA	ATGTCCTCCT	GGAAGATGGC	780
	AAGGGCGGTT	GTCCCTTCGA	CCCGAATTTC	AAGTCCACTG	CCTGTTGGT	TGATGGCGAG	840
	CTCTACACTG	GAACAGTCAG	CAGCTTCCAA	GGGAATGACC	CGGCCATCTC	GCGGAGCCAA	900
	AGCCTTGCGC	CCACCAAGAC	CGAGAGCTCC	CTCAACTGGC	TGCAAGACCC	AGCTTTTGTG	960
75	GCTCAGCCT	ACATTCTCTG	GAGCCTGGGC	AGCTTGCAAG	GCGATGATGA	CAAGATCTAC	1020
	TTTTTCTTCA	GCGAGACTGG	CCAGGAATTT	GAGTCTCTTG	AGAACACCAT	TGTGTCCCGC	1080
	ATTGCCCCGA	TCTGCAAGGG	CGATGAGGT	GGAGAGCGGG	TGCTACAGCA	GCGCTGGACC	1140
	TCCTTCTCTA	AGGCCAGCT	GCTGTGCTCA	CGGCCCGAGC	ATGGCTTCCC	CTTCAACGTG	1200
	CTGCAGGATG	TCTTCAAGCT	GAGCCCCAGC	CCCCAGGACT	GGCGTGACAC	CCTTTTCTAT	1260
80	GGGGTCTTCA	CTCTCCAGTG	GCAAGGGGA	ACTACAGAAG	GCTCTGCGGT	CTGTGTCTTC	1320
	ACAATGAAGG	ATGTGCAGAG	AGTCTTCAGC	GGCCTCTACA	AGGAGGTGAA	CGGTGAGACA	1380
	CAGCAGTGGT	ACACCGTGAC	CCACCGGTG	CCCAACCCCG	GGCCTGGAGC	GTGCATCACC	1440
	AACAGTCCCC	GGGAAAGGAA	GATCAACTCA	TCCCTGCAGC	TCCCGAGCCG	CGTGCTGAAC	1500
	TTCTCAAGG	ACCACTTCTC	GATGGAGCGG	CAGGTCCGAA	GCCGATGCT	GCTGTGACAC	1560
85	CCCCAGGCTG	GCTACCAAGG	CGTGGCTGTA	CACCGCGTCC	CTGGCCTGCA	CCACACTTAC	1620
	GATGTCTCTT	TCCTGGGCGC	TGGTGACGGC	CGGCTCCACA	AGGCAGTGAG	CGTGGGCCCC	1680
	CGGGTGACA	TCATTGAGGA	GCTGCAGATC	TCTCATCGG	GACAGCCCGT	GCAGAACTCT	1740

5
10
15
20
25
30
35

CTCCTGGACA CCCACAGGGG GCTGCTGTAT GCGGCCTCAC ACTCGGGCGT AGTCCAGGTG 1800
CCCATGGCCA ACTGCAGGCT GTACAGGAGC TGTGGGGAAT GCCTCCTCGC CGGGAACCCC 1860
TACTGTGCTT GGAGCGGCTC CAGCTGCAAG CAOGTCAGCC TCTACCAGCC TCAGCTGGCC 1920
ACCAGGCGCT GGATCCAGGA CATCGAGGA GCCAGCGCCA AGGACCTTTG CAGCGCTCT 1980
TCGGTTGTGT CCCGCTCTT TGTACCAACA GGGGAGAAGC CATGTGAGCA AGTCCAGTTC 2040
CAGCCCAACA CAGTGAACAC TTTGGCCTGC CGCTCTCTCT CCAACTTGGC GACCCGACTC 2100
TGGCTACGCA ACGGGGCCCC CGTCAATGCC TCGGCTCTCT GCCAGTGTCT ACCCACTGGG 2160
GACCTGCTGC TGGTGGGCAC CCAACAGCTG GGGGAGTTCC AGTGTCTGTC ACTAGAGGAG 2220
GGCTTCCAGC AGCTGGTAGC CAGCTACTGC CCAGAGGTGG TGGAGGACGG GGTGGCAGAC 2280
CAACAGATG AGGGTGGCAG TGTACCGGTC ATTATCAGCA CATCGCGTGT GAGTGCACCA 2340
GCTGGTGGCA AGGCGAGCTG GGGTGCAGAC AGGTCTTACT GGAAGGAGTT CTTGGTGTATG 2400
TGCACGCTCT TTTGCTGGC CGTGTGCTC CCAGTTTTAT TCTTGTCTTA CCGGCACCGG 2460
AACAGCATGA AAGTCTTCTT GAAGCAGGGG GAATGTGCCA GGTGTGACCC CAAGACCTGC 2520
CCTGTGGTGC TGCCCCCTGA GACCCGCCCA CTCAACGCCC TAGGGCCCCC TAGCACCCCG 2580
CTCGATCACC GAGGTACCA GTCCCTGTCA GACAGCCCCC CGGGTCCCCG AGTCTTCACT 2640
GAGTGCAGCA AGGAGGCACT CAGCATCCAA GACAGCTTCG TGGAGGTATC CCCAGTGTGC 2700
CCCGCGCCCC GGGTCCGCTT TGGCTCGGAG ATCGGTGACT CTGTGGTGTG AGAGCTGACT 2760
TCCAGAGGAC GCTGCCCTGG CTTGAGGGGC TGTGAATGCT CGGAGAGGGT CAACTGGACC 2820
TCCCTCCGCG TCTGCTCTTC GTGGAACACG ACCGTGGTGC CCGGCCCTTG GGAGCCTTGG 2880
GGCCAGCTGC CCGTCTGCTC TCCAGTCAAG TAGCGAAGCT CCTACCACC AGACCCCAA 2940
ACAGCGGTGG CCCAGAGGCT CCGTGGCCAA TATGGGGGCC TGCTAGGTT GGTGGAAACG 3000
TGCTCTTAT GTAACTGAG CCCTTTGTTT AAAAAACAAT TCCAAATGTG AACTAGAAAT 3060
GAGAGGGAAG AGATAGCATG GCATGCAGCA CACACGGCTG CTCAAGTTC TGGCTCCCA 3120
GGGTGTCTGG GGATGCATCC AAAGTGGTTG TGTGAGACAG AGTTGGAAC CCTCACCAAC 3180
TGGCTCTTTC ACCTTCCACA TTATCCGCT GCCACCGCT GCCCTGTCTC ACTGCAGATT 3240
CAGGACCAAG TTGGGCTGCG TCGTCTGCTC CTTGCCAGTC AGCCGAGGAT GTAGTTGTG 3300
CTGCGCTGCT GACCAACCT CAGGAGCAG AGGGCTAGGT TGGCACTCG GCCCTACCA 3360
GGTCTGGGC TCGGACCCAA CTCTGGACC TTTCCAGCT GTATCAGCT GTGCTCACAC 3420
GAGAGGACAG CGCGAGCTCA GGAGAGATTT CGTGACAATG TAGCCTTTC CCTCAGAATT 3480
CAGGGAAGAG ACTGTGCGCT GCCTTCTTCC GTTGTGCGT GAGAACCGT GTGCCCTTTC 3540
CCACCATATC CACCTCGCT CCATCTTGA ACTCAAACAG GAGGAACATA CTGACCCCTG 3600
GTCCTCTCCC CAGTCCCCAG TTCACCTCC ATCCCTCAC TTCTTCCAT CTAAGGGATA 3660
TCAACACTGC CCAGCAGAGG GGCCCTGAAT TTATGTGGTT TTTATACATT TTTTAAATAG 3720
ATGCATTTA TGTATTTTT TAATAAAGTC TGAAGAATTA CTGTTT

Seq ID NO: 479 Protein sequence
Protein Accession #: XP_044533.3

40
45
50
55

1 11 21 31 41 51
MLRTAMGLRS WLAAPWGALP PRPPLLLLLL LLLLLQPPPP TWALSPRISL PLGSEERPFL 60
RFEAEHISNY TALLLSRDGR TLYVGAREAL FALSNLSFL PGGEYQELLW GADAEKKQQC 120
SPKGDQPRD QONYIKILLP LSGSHLPTCG TAAPSPMCTY INMENFTLAR DEKENVLLED 180
GKGRCPDPFN FKSTALVVDG ELYTGTVSSF QGNDFPAIRS QSLRPTKTES SLNWLQDPAP 240
VASAYIPESL GSLQDDDKI YFFSETGQE PEFENTIVS RIARICKGDE GGERVLQQRW 300
TSFLKAQLLC SRPDDGFPPN VLQDVFTLSP SPQDWRDTLP YGVFTSQWER GTTEGSAVCV 360
FTMKDVQRFV SGLYKEVNR TQWYTVTHP VPTPRPGACI TNSARERKIN SSLQLPDRVL 420
NFLKDHFLMD QVRSRMLLL QPARYQRYA VHRVPLHHT YDVLFLGTGD GRLEKAVSVG 480
PRVHIIIEELQ IFSSGQPVQN LLLDTERGLL YAASHSGVVQ VPMANCSLYR SCGDCLLARD 540
PYCAWSGSSC KHVSLYQPL ATRPWIQDIE GASAKDLCSA SSVSPSPFPV TGEKPCBQVQ 600
PQPNTVMTLA CPLLNLNATR LNLRNAPVN ASASCHVLEP GDLLLVTGQQ LGEFQCWSLE 660
EGPQLVASY CPEVEDGVA DQDEGGSVV VIISTSRVSA PAGGKASHGA DRSYWKELFV 720
MCTLFVLAVL LPLVFLYRHR RNSMKVFLKQ GECASVHPKT CPVLLPETR PLNGLGPPST 780
PLDHRGQSL SDSPPGSRVF TESEKRPLSI QDSFVSVSPV CPRPRVRLGS EIRDSV

Seq ID NO: 480 DNA sequence
Nucleic Acid Accession #: NM_004217.1
Coding sequence: 58..1092

60
65
70
75
80
85

1 11 21 31 41 51
GGCCGGGAGA GTAGCAGTGC CTTGGACCCC AGCTCTCCTC CCCCTTTCTC TCTAAGGATG 60
GCCAGAAAG AGAACTCCTA CCCTGGCCCC TACGGCGGAC AGACGGCTCC ATCTGGCCTG 120
AGCACCTGCG CCCAGCGAGT CTTCCGGAAA GAGCCTGTCA CCCCATCTGC ACTTGTCTCTC 180
ATGAGCGGCT CCAATGTCCA GCCACAGCT GCCCTGGGCC AGAAGGTGAT GGAGAATAGC 240
AGTGGGACAC CGACATCTT AACGCGGCAC TTCACAATTG ATGACTTTGA GATTGGGCGT 300
CCTCTGGGCA AAGGCAAGTT TGGAAACGTG TACTTGGCTC GGGAGAAGAA AAGCCATTTC 360
ATCGTGGGCG TCAAGGTCTT CTTCAAGTCC CAGATAGAGA AGGAGGGCGT GGAGCATCAG 420
CTGCGCAGAG AGATCGAAAT CCAGGCCAC CTGCAACATC CCAACATCTC CGCTCTCTAC 480
AACTATTTTT ATGACCGGAG GAGGATCTAC TTGATTCTAG AGTATGCCCC CCGCGGGGAG 540
CTCTACAAGG AGCTGCAGAA GAGCTGCACA TTTGACGAGC AGCGAACAGC CACGATCATG 600
GAGGAGTTGG CAGATGCTCT AATGTACTGC CATGGGAAGA AGGTGATTCA CAGAGACATA 660
AAGCCAGAAA ATCTGCTCTT AGGGCTCAAG GGAGAGCTGA AGATTGCTGA CTTCCGCTGG 720
TCTGTGCATG CGCCTCTCTT GAGGAGGAAG ACAATGTGTG GCACCTTGGA CTACCTGCC 780
CCAGAGATGA TTGAGGGGCG CATGCACAAT GAGAAGGTGG ATCTGTGGTG CATTGGAGTG 840
CTTTGCTATG AGTGCTGGT GGGGAACCCA CCCTTTGAGA GTGCATCACA CAACGAGACC 900
TATCGCGGCA TCGTCAAGGT GAGCTTAAAG TTCCCGCTT CTGTGCCCAC GGGAGCCCAG 960
GACCTCATCT CCAAACCTCT CAGGCATAAC CCCTCGGAAC GGCTGCCCTT GGCCTAGGTC 1020
TCAGCCACCC CTTGGGTCCG GGCCAACTCT CGGAGGGTGC TGCTCCCTC TGCCCTTCAA 1080
TCTGTGCGCT GATGTCCCT GTCACTACT CGGGTGCCTG TGTGTGTATG TCTGTGTATG 1140
TATAGGGGAA AGAAGGGATC CCTAAGTGT CCCTTATCTG TTTTCTACCT CCTCTTTGT 1200
TTAATAAAG CTGAAGCTTT TTGT

Seq ID NO: 481 Protein sequence
Protein Accession #: NP_004208

1 11 21 31 41 51

MAQKENSYPW PYGRQTPSPG LSTLPQVRVL KEFVTPSALV LMSRSNVQPT AAPGQKVMEN 60
SSGTPDILTR HFTIDDFEIG RPLGKGFNG VYLAREKKSH FIVALKVLPK SQIEKEGVEH 120
QLRRREIEQA HLHHPNILEL YNYFYDRRLI YLILEYAPRG ELYKELQKSC TFDQORTATI 180
MEELADALMY CHGKQVIHRD IKPENLLGL KGELKIADFG WSVHAPSLRR KTMCGTLDYL 240
PPEMIEGRMH NEKVDLWCIG VLYCELLVGN PPFESASHNE TYRRIVKVDL KFPASVPTGA 300
QDLISKLLRH NPSERLPLAQ VSAHPWVRAN SRRVLPPSAL QSVN

Seq ID NO: 482 DNA sequence
Nucleic Acid Accession #: AK055663
Coding sequence: 38..1423

1 11 21 31 41 51
15 AGAAGCGCTT CCGGCGGGAG CTGTGCAGCT CCTTATCATG GGGACAATTC ATCTCTTTCTG 60
AAAACACAAA AGATCTCTTT TTGGCAAGTT GTTACGGGAA TTATAGACTG TAGCAGCTGA 120
CCGAAGGTCCT TGAAGATAC TGCTCTTTGG TGTAATAAAC TTGATATGTA CTGGCTTCTT 180
GCTTATGTGG TGCAGTTCTA CTAATAGTAT AGCTTTAACT GCCTATACTT ACCTGACCAT 240
TTTTGATCTT TTTAGTTTAA TGACATGTTT AATAAGTTAC TGGGTAACAT TGAGGAAACC 300
20 TAGCCCTGTG TATTCATTTG GGTGTGAAAG ATTAGAAGTC CTGGCTGTAT TTGCCTCCAC 360
AGTCTTGGCA CAGTGGGAG CTCTCTTTAT ATTAAGAAAG AGTGCAGAAC GCTTTTGGGA 420
ACAGCCCGAG ATACAGACGG GAAGATTATT AGTTGGTACT TTTGTGGCTC TTTGTTTCAA 480
CCTGTTCAGG ATGCTTCTTA TTGGGAATAA ACCTTTTGCT TATGTCTCAG AAGCTGCTAG 540
TACGAGCTGG CTTCAGAGC ATGTTGCAGA TCTTAGTCTG AGCTTGTGTG GAATTATTCC 600
25 GGGACTTAGC AGTATCTTCC TTCCCGAAT GAATCCATTG GTTTTGATTG ATCTTGCTGG 660
AGCATTGTCT CTCTGTATTA CATATATGCT CATTTGAAAT AATAATTATT TTGCGGTAGA 720
CACTGCTCTT GCTATAGCTA TTGCCCTGAT GACATTGGCC ACTATGTATC CCATGAGTGT 780
GTACAGTGGG AAGTCTTAC TCCAGACAA ACCACCCCAT GTTATTGGTC AGTTGGACAA 840
30 ACTCATCAGA GAGGTATCTA CCTTAGATGG AGTTTATGAA GTCCGAAATG AACATTTTGT 900
GACCCTAGT TTTGGCTCAT TGGCTGGATC AGTGCATGTA AGAATTGCAG GAGATGCCAA 960
TGAACAAATG GTTCTGTCTC ATGTGACCAA CAGGCTGTAC ACTCTAGTGT CTACTCTAAC 1020
TGTTCAAAT TTTCAAGGATG ACTGGATTAG GCCTGCTCTA TTGCTGGGCG CTGTTGCAGC 1080
CAATGTCTTA AACTTTTCTG ATCATCACTT AATCCCAATG CCTCTTTTAA AGGGTACTGA 1140
35 TGATTTGAAC CCACTTACAT CACTCCAGC TAAACCTAGT AGTCCACCTC CAGAATTTTC 1200
ATTTAACTCA CCGGGGAAA ATGTGAACCC AGTTATTCTT CTAAACACAC AAACAAGGCC 1260
TTATGTTTTT GGTCTCAATC ATGACACAC ACCTTACAGC AGCATGCTTA ATCAAGGACT 1320
TGAGTTTCCA GGAATGGAG CACTCAAGG ATTGAGGACT GGTTTTACAA ATATACCAAG 1380
TAGATATGGA ACTAATAATA GAATTGACA ACCAAGACCA TGATAGACTC TAACCTATT 1440
40 TTATAAGGAA TATGACTCC TTGGCTTCCA ATTTATTAG TAATCCAACT TTGCATTGAC 1500
TGTTTAATCA CCGGCTTCAA ATGTTAGATA ATAGTAGTCT TGTTCACTT TCATGAAACC 1560
TATGAACTTA TATTTTGTGA AAATGTATTT GTGACAGTGA AATCCTCGTA AATGTTAAAG 1620
GCITTAATAA GGCTCTCTTT AGAAAATGTG TTTCTTAAAT TTTGGAATTT GGTATCTTTG 1680
GTTTTGTAGT TGACTGCACT GTGATGTGAC CTTACCTTTA TAAGAGCCAC TTGATGGAGT 1740
45 AGATCTGTCA CATTACTAAG ATACGATATT TCTTTTTTTT TCCGAGACGG AGTCTTGCTC 1800
TGCCACTGTG CCGGCGCAAT ACATTATTAT TAACTTAAGG CTGTACTTTA TTAAGGCTTC 1860
CTTAGTTTTT GTTTTGTGTT GTTTTGTGAG ATGGAGTCTC ACTCTGTCGC CCAGGCTGGA 1920
ATGCACTGGC ATGATCTCAG CTCACCTGCA CCTCTGCCTC CTGAGTTCAA ATGATTCTCC 1980
TGCCCTCAGC TCCCGAGTAG CTGGGATTAC AGGCACCTGC CACCACGCCC AGCTAATTTT 2040
50 TGTATTTTGA GTAAAGACCG GGGATTTTAC CATGTTGGCC AGGCTGGTCT TGAACCTCTG 2100
ACCTCATGAT CCACTCACTT TAGCCTCCCA AAGTGTGGG ATTAGGTGTG AGCCACCGCA 2160
CCTGGCCGAT ATTTTCTTTA ATGAAATTTA TAAATATGCT TCTTGAATAA TACACATTTT 2220
GGGAAAGGGA AAATGTCTTG TTCAAAAAGT AAAGGTCTCT TTTATAGCTT TTCCAACTT 2280
AATTGTAAA TTTTCTTTG AGGTTCTCCT GAATTATGTC TTACAACTA AAAGCAAAA 2340
TTTTTAGCAG AAATTTTGA ATACATCTTA TCTAGCACAA TTTGAATTT TAATTATCAA 2400
55 GATTTTGTG AAAGTTTCTC TCCTTTAAAA ATTTTAGTAC ATTTGTAAT

Seq ID NO: 483 Protein sequence
Protein Accession #: BAB70980.1

1 11 21 31 41 51
60 MGIHLFRKP QRSFFGKLLR EPRIVAADR SWKILLFGVI NLICTGFLLM WCSSTNSIAL 60
TAYTYLTIFD LPSLMTCLIS YWVTLRKPS VYSFGFERLE VLAVFASTVL AQLGALFILK 120
65 ESAERFLEQ BIHTGRLLVG TFVALCFNLF TMSIRNKPF AYVSEAASTS WLQEHVADLS 180
RSLCGIIPGL SSIFLPRMNP FVLIDLAGAF ALCITYMLIE INNYFAVDTA SAIAIALMTF 240
GTMYPMVSYS GKVLQTPPP HVIGQLDKLI REVSTLDGVL EVRNEHFWTL GFGSLAGSVH 300
VRIRRDANEQ MVLAEVNTNL YTLVSTLTQV IPKDDWIRPA LLSGPVAANV LNFSDHNVIP 360
MPLLKGTDDL NPVTSTPAKP SSPPEFSFN TPGKNVNFVI LLNTQTRPYG FGLNHGHTPY 420
70 SSMNLGGLGV FGIATQGLR TGFTNIPSRY GTNNRIGQPR P

Seq ID NO: 484 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..900

1 11 21 31 41 51
75 ATGCCGCCGC GGGAGCTGAG CGAGGCGGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC 60
CCGCGGCGGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGTGGT GGGCGACGGC 120
80 CCGGTGGGCA AGAGAGCCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGGCTTAC 180
CGGCCCACTG CGCTGGACAC CTTCTCTGGT ACGTACGTTT AATCGCCCGT GGGCGCGCGT 240
GGCTGCGGCG GGGCTGTGCA CCGGGAGCT GGGCGCGGCG TCTCGGCGGG AGGGCGCAGA 300
GGACCCCGGG GAGGAGACTG GAGCAGGCCC CGAGGTGGCG CTGGTGGGCT CCAGGACGCT 360
CTTCTAACT CAGGCTCTCC CCGCCCGGCC CCGCAGTGC AAGTCTGGT GAGTGGAGCT 420
85 CCGGTGCGCA TTGAGCTCTG GACACAGCG GACAGGAGG ATTTGACCG ACTTCGTTC 480
CTTTGCTACC CGGATACGA TGTCTTCTG GCGTGTCTTA GCGTGGTGA GCCCAGCTCC 540
TTTCAAAACA TCACAGAGAA ATGCTGCCC GAGATCCGCA CGCACAAACC CCAGGCGCCT 600
GTGCTGCTGG TGGGACCCCA GGCGGACCTG AGGGACGATG TCAACGTACT AATTGAGCTG 660

GACCAAGGGG GCGGGGAGGG CCCCGTGGCC CAACCCAGG CTCAGGGTCT GCGCGAGAAG 720
 ATCCGAGCCT GCTGCTACCT TGAGTGCTCA GCCTTGACGC AGAAGAATT GAAGGAAGTA 780
 TTTGACTCGG CTATTCTCAG TGCCATTGAG CACAAAGCCC GGCTGGAGAA GAAACTGAAT 840
 GCCAAGGTG TCGCACCCCT CTCGCCCTGC CGCTGGAAGA AGTTCCTCTG CTTGCTTTGA

Seq ID NO: 485 Protein sequence
 Protein Accession #: PGENESH predicted

1 11 21 31 41 51
 MPPRELSEAE PPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60
 RPTALDTFSG TYVQSPVRRP GCGGAVERGA GAGVSAGGR GPRGGDWSRP RGGAGAAQDA 120
 LPNSGSPRPA PAVQVLVDGA PVRIELWDTA QGEDFDRLRS LCYPDTDVFL ACPVSVVQPS 180
 FQNIKEWLP EIRTHNPQAP VLLVGTQADL RDDVNVLIQL DQGGREGPVP QPQAGLAER 240
 IRACCYLECS ALTQKNLKEV FDSAILSIAE HKARLEKKLN AKGVRTLSRC RWWKPPCFV

Seq ID NO: 486 DNA sequence
 Nucleic Acid Accession #: XM_063832.2
 Coding sequence: 1..711

1 11 21 31 41 51
 ATGCCGCGCG GGGAGCTGAG CGAGGCGGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC 60
 CCGCGCGCGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGTCTGT GGGCGACGGC 120
 GCGGTGGGCA AGAGCAGCCT CATGTCAGC TACACCTGCA ATGGGTATCC CGCGCGCTAC 180
 CGGCCCACTG CGCTGGACAC CTTCTCTGTG CAAGTCTCTG TGGATGGAGC TCCGGTGGCG 240
 ATTGAGCTCT GGGACACAGC GGGACAGGAG GATTTTGACC GACTTCGTTT CTTTGTCTAC 300
 CCGGATACCG ATGTCTCTCT GCGGTGCTTC AGCGTGTGTC AGCCCACTC CTTTCAAAC 360
 ATCAGAGAGA AATGSGCTGC CGAGATCCGC ACGCACAAAC CCCAGGCGCC TGTGCTGCTG 420
 GTGGGCACCC AGGCCGACCT GAGGACGAT GTCAACGTAC TAATTCAGT GAGCAGGGG 480
 GGGCGGAGG GCGCCGTGCC CCAACCCAG GCTCAGGTC TGGCCGAGAA GATCCGAGCC 540
 TGCTGTCTAC TTGAGTGCTC AGCCTTGAGC CAGAAGAACT TGAAGGAAGT ATTTGACTCG 600
 GCTATTCTCA GTGCCATTGA GCACAAAGCC CGGCTGGAGA AGAACTGAA TGCCAAAGGT 660
 GTGCGCACCC TCTCCGCTG CCGCTGGAAG AAGTCTCTCT GCTTCGTTTG A

Seq ID NO: 487 Protein sequence
 Protein Accession #: XP_063832.1

1 11 21 31 41 51
 MPPRELSEAE PPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60
 RPTALDTFSG QVLVDGAPVR IELWDTAGQE DFDRLRSLCY PDTDVFLACF SVVQPSFQ 120
 ITEKWLPEIR THNPQAPVLL VGTQADLRDD VNVLIQLDQG GREGPVPOPO AQGLAEKIRA 180
 CCYLECSALT QKNLKEVPDS AILSAIEHKA RLEKKLNAGK VRTLSRCRWK KFFCFV

Seq ID NO: 488 DNA sequence
 Nucleic Acid Accession #: NM_014398.1
 Coding sequence: 64..1314

1 11 21 31 41 51
 GGCACCGATT CGGGGCTGCG CCGGACTTCG CCGCACGCTG CAGAACCTCG CCCAGCGCCC 60
 ACCATGCCCC GGCAGCTCAG CGCGGCGGCC GCGCTCTTCG CGTCCCTGGC CGTAATTTTG 120
 CACGATGGCA GTCAAAATGAG AGCAAAAGCA TTTCCAGAAA CCAGAGATTA TTCTCAACCT 180
 ACTGACAGAG CAACAGTACA GGACATAAAA AAACCTGTCC AGCAACAGAG TAAGCAAGCA 240
 CCTCACAAA CTTTAGCAGC AAGATTATG GATGGTCATA TCACCTTTCA AACAGCGGCC 300
 ACAGTAAAA TTCCAACAAC TACCCAGCA ACTACAAAA ACACCTGCAAC CACAGCCCA 360
 ATTACCTACA CCGTGTCTAC AACCCAGGCC ACACCAACA ACTCACACAC AGCTCCTCCA 420
 GTTACTGAAG TTACAGTCCG CCCTAGCTTA GCCCTTATT CACTGCCACC CACCATCACC 480
 CCACAGCTC ATACAGCTCG AACCACTTCA TCAACCGTCA GCCACACAAC TGGGAACACC 540
 ACTCAACCA GTAAACAGAG CACCTTCCA GCAACTTTAT CGATAGCACT GCACAAAGC 600
 ACAACCGGTG AGAAGCCTGA TCAACCCACC CATGCCCCAG GAACACGGC AGCTGCCACC 660
 AATAACACCC GCACAGCTGC ACCTGCCTCC ACGGTTCTCG GGGCCACCCT TGCACCTCAG 720
 CCATCGTCAG TCAAGACTCG AATTATCAG GTTCTAAAG GAAGCAGACT CTGTATAAAA 780
 CGAGAGATGG GGATACAGCT GATTGTTCAA GACAAGGAGT CGGTTTTTTC ACCTCGGAGA 840
 TACTTCAACA TCGACCCCAA CGCAACGCAA GCGTCTGGGA ACTGTGGCAC CGAAAAATCC 900
 AACCTTCTGT TGAATTTTCA GGGCGGATTT GTGAATCTCA CATTTACCA GGATGAAGAA 960
 TCATATTATA TCAGTGAAGT GGGAGCCTAT TTGACCGTCT CAGATCCAGA GACAGTTTAC 1020
 CAAGGAATCA AACATGCGGT GGTGATGTT CAGACAGCAG TCGGCAATTC CTTCAAGTGC 1080
 GTGAGTGAAC AGAGCCTCCA GTTGTGAGCC CACCTGCAGG TGAACAACA CGATGTCCAA 1140
 CTTCAAGCCT TTGATTTTGA AGATGACCAC TTTGGAATG TGGATGAGTG CTGCTGTGAC 1200
 TACACAATTG TGCTTCTGT GATTGGGGCC ATCGTGGTTG GTCTCTGCCT TATGGSTATG 1260
 GGTGTCTATA AAATCCGCTC AAGGTGTCAA TCATCTGAGT ACCAGAGAA CTAAATTTGT 1320
 CCGGGGGGGA ATGAAAATAA TGAATTTAG AGAACTCTTT CATCCCTTCC AGGATGGATG 1380
 TTGGGAATTT CCCTCAGAGT GTGGGTCTTT CAAACAATGT AAACCAACAT CTTCTATTCA 1440
 AATGAAGTGA GTCATGTGTG ATTTAAGTTC AGGCAGCACA TCAATTTCTA AATACTTTTT 1500
 GTTATTTTAA TGAAGATAT AGTGAGCTGT TTATTTTCTA GTTTCCTTTA GAATATTTTA 1560
 GCCACTCAAA GTCAACATT GTGATATGTT GAATTAACAT AATATATGTA AAGTAGAATA 1620
 AGCCTTCAAA TTATAAACA AGGGTCAATT GTAACATAA CTACTGTGTG TGCAATTGAAG 1680
 ATTTTATTTT ACCCTTGATC TTAACAAAGC CTTTGTCTTG TTATCAAAAT GACTTTCACT 1740
 GCTTTTACTA TCTGTGTTTT ATGGTTTCAT GTAACATACA TATTCCTGGT GTAGCACTTA 1800
 ACTCCTTTTC CACTTTAAAT TTGTTTTTGT TTTTGGAGAC GGAGTTTCAC TCTTGTCAAC 1860
 CAGGCTGGAG TACAGTGGCA CGATCTCGGC TTATGGCAAC CTCGCGCTCC CGGGTTCAAG 1920
 TGATTTCTCT GCTTCAGCTT CCGGAGTAGC TGGGATTACA GGCACACACT ACCAGCCTG 1980
 GCTAATTTT GTATTTTAT TATAGACGGG TTTCAACATG TTGGCCAGAC TGGTCTTGAA 2040
 CTCCTGACCT CAGGTGATCC ACCCACTTCA GCCTCCCAA GTGCTGGGAT TACAGGCATG 2100
 AGCCATTGCG CCGGCTCTTA AATGTTTTTT TTAATCATCA AAAAGAACAA CATATCTCAG 2160

5
10
15

```

GTTGTCTAAG TGTTTTATG TAAACCAAC AAAAGAACA AATCAGCTTA TATTTTTTAT 2220
CTTGATGACT CCTGCTCCAG AATTGCTAGA CTAAGAATTA GGTGGCTACA GATGGTAGAA 2280
CTAAACAATA AGCAAGAGAC AATAATAATG GCCCTTAATT ATTAACAAGG TGCCAGAGTC 2340
TAGGCTAAGC ACTTTATCTA TATCTCATTT CATCTCACA ACTTATAAGT GAATGAGTAA 2400
ACTGAGACTT AAGGGAACCTG AATCACTTAA ATGTACACCTG GCTAACTGAT GGCAGAGCCA 2460
GAGCTTGAAT TCATGTTGGT CTGACATCAA GGTCTTTGGT CTTCTCCCTA CACCAAGTTA 2520
CCTACAAGAA CAATGACACC ACACCTCTGCC TGAAGGCTCA CACCTCATAC CAGCATACGC 2580
TCACCTTACA GGGAAATGGG TTTATCCAGG ATCATGAGAC ATTAGGGTAG ATGAAAGGAG 2640
AGCTTTGCAG ATAACAAAAT AGCCTATCCT TAATAAATCC TCCACTCTCT GGAAGGAGAC 2700
TGAGGGGCTT TGTAAACAT TAGTCAGTTG CTCAATTTTA TGGGATTGCT TAGCTGGGCT 2760
GTAAGATGA AGGCATCAA TAAACTCAA GTATTTTAA ATTTTGTGA TAATAGAGAA 2820
ACTTCGCTAA CCACTGTTC TTTCTTGAGT GTATAGCCCC ATCTTGTGCT AACTTGTCTG 2880
TTCTGCATT CATATCCATA TTTCTATTG TTCACTTTAT TCTGTAGAGC AGCCTGCCAA 2940
GAATTTTATT TCTGCTGTTT TTTTGTGTC TAAAGAAAGG AACTAAGTCA GGATGTTAAC 3000
AGAAAAGTCC ACATAACCTT AGAATTTCTA GTCAAGGAAT AATTCAAGTC AGCCTAGAGA 3060
CCATGTTGAT TTCTCTCATG TGTTCCTTGA TGACTCAGTA AGTTGGCAAG GTCTGACTTT 3120
TAGTCTTAAT AAAACATGTA ATTGTAGTAA AGGTTTTTGC AATAAAAACT TACTTTGG

```

Seq ID NO: 489 Protein sequence
Protein Accession #: NP_055213.1

20
25
30

```

1 11 21 31 41 51
MPRQLSAAA LFASLAVILH DGSQMRKAP PETRDYSQPT AAATVQDIKK PVQPAKQAP 60
HQTLLAEMFD GHITPQTAA VKIPTTTPAT TKNTATTSP I TYTLVTQTAT PNNSHAPPV 120
TEVTVGPSLA PYSLPPTITP PAHTAGTSSS TVSHTTGNTT QPSNQTTLPA TLSIALHKST 180
TQKFPDQPTH APGTTAAAHN TTRTAAPAST VPGFTLAPQP SSVKTGIYQV LINGSRLCIA 240
EMGIQLVQD KESVSPRRY FNIDPNATQA SGNCGTRKSN LLLNFQGGFV NLFTTDEES 300
YVISEVGAYL TVSDPETVYQ GIKHVVVMFQ TAVGHSFKCV SEQSLLSAH LQVKTIDVQL 360
QAFDFEDDHF GNVDCSSDY TIVLPVIGAI VWGLCLMGNG VYKIRLRQCS SGYQRI

```

Seq ID NO: 490 DNA sequence
Nucleic Acid Accession #: NM_005409.3
Coding sequence: 94..378

35
40
45
50
55
60

```

1 11 21 31 41 51
TTCTTTTCAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAAGCT GAAGTAGCAG 60
CAACAGCACC AGCAGCAACA GCAAAAAACA AACATGAGTG TGAAGGGCAT GGCTATAGCC 120
TTGGCTGTGA TATTGTGTGC TACAGTTGTT CAAGGCTTCC CCATGTTCAA AAGAGGACGC 180
TGTCTTTGCA TAGGCCCTGG GGTAAAAGCA GTGAAAGTGG CAGATATTGA GAAAGCCTCC 240
ATAATGTACC CAAGTAACAA CTGTGACAAA ATAGAAGTGA TTATTACCTT GAAAGAAAAT 300
AAGAGACAAG GATGCTTAAA TCCCAAATCG AAGCAAGCAA GGCTTATAAT CAAAAGATT 360
AAGAGAAAGA ATTTTAAAA ATATCAAAAC ATATGAAGTC CTGGAAGAG GCATCTGAAA 420
AACCTAGAAC AAGTTAACT GTGACTACTG AAATGACAAG AATTCTACAG TAGGAAACTG 480
AGACTTTTCT ATGGTTTGTG GACTTTTCAAC TTTTGTACAG TTATGTGAAG GATGAAAGGT 540
GGGTGAAAGG ACCAAAAACA GAAATACAGT CTTCCTGAAT GAATGACAAT CAGAATTCCA 600
CTGCCAAAG GAGTCCAGCA ATTAATAGGA TTTCTAGGAA AAGCTACCTT AAGAAAGGCT 660
GGTTACCATC GGAGTTTACA AAGTGCTTTC ACGTTCTTAC TTGTTGTATT ATACATTCT 720
GCATTCTCAG CTAGAGAAC CTCTAGATT TGATGCTTAC AACTATTCTG TTGTGACTAT 780
GAGAACATTT CTGTCTCTAG AAGTTATCTG TCTGTATTGA TCTTTATGCT ATATTACTAT 840
CTGTGGTTAC AGTGGAGACA TTGACATTAT TACTGGAGTC AAGCCCTTAT AAGTCAAAAG 900
CATCTATGTG TGTAAAGCA TTCTCAAAAC ATTTTTCAT GCAATACAC ACTTCTTTCC 960
CCAAATATCA TGTAGCACAT CAATATGTAG GGAACATTC TTATGCATCA TTTGGTTTGT 1020
TTTATAACCA ATTCAATTAA TGTAAATCAT AAAATGTACT ATGAAAAAAA TTATACGCTA 1080
TGGGATACCT GCAACAGTGC ACATATTCCA TAACCAAAAT AGCAGCACCG GTCTTAATTT 1140
GATGTTTTTC AACTTTTATT CATGAGATG TTTTGAAGCA ATTAGGATAT GTGTGTTTAC 1200
TGTACTTTTT GTTTGTATCC GTTTGTATAA ATGATAGCAA TATCTGGAC ACATTGAAA 1260
TACAAAATGT TTTTGTCTAC CAAAGAAAAA TGTGAAAAA TAAGCAAATG TATACCTAGC 1320
AATCACTTTT ACTTTTGTGA ATTCTGTCTC TTAGAAAAAT ACATAATCTA ATCAATTTCT 1380
TTGTTATGTC CTATATACCT TAAATTTAG GTATACTCAA GACTAGTTTA AAGATCAAA 1440
GTCAATTTTT TCTTAATAA ACTACACAA CCTTCTTTT TTAACAAAAA AAA

```

Seq ID NO: 491 Protein sequence
Protein Accession #: NP_005400.1

65
70

```

1 11 21 31 41 51
MSVKGMAIAL AVILCATVVQ GPFMPKRGRC LCIGPGVKAV KVAIDIEKASI MYPNNCDKI 60
EVIITLKENK GQRCLNPKSK QARLIKKVE RKNF

```

Seq ID NO: 492 DNA sequence
Nucleic Acid Accession #: NM_000577.1
Coding sequence: 41..520

75
80
85

```

1 11 21 31 41 51
GGCAGGAGGG GAAGACCTCC TGTCTTATCA GGCCCTCCCC ATGGCTTTAG AGACGATCTG 60
CGGACCTCTT GGGAGAAAAT CCAGCAAGAT GCAAGCCTTC AGAATCTGGG ATGTTAAACA 120
GAAGACCTTC TATCTGAGGA ACAACCAACT AGTTGCCGGA TACTTGCAAG GACCAATATG 180
CAATTTAGAA GAAAAGATAG ATGTGGTACC CATTGAGCCT CATGCTCTGT TCTTGGGAAT 240
CCATGGAGGG AAGATGTGCT GTCTCTGTGT CAAGTCTGGT GATGAGACCA GACTCCAGCT 300
GGAGGCAGTT AACATCACTG ACCTGAGCGA GAACAGAAAG CAGGACAAGC GCTTCGCCTT 360
CATCCGCTCA GACAGTGGCC CCACCACAGG TTTTGAAGT CCGCCTGCC CCGGTTGGTT 420
CCTCTGCAAC GCGATGGGAG CTGACCAGCC CGTCAGCCTC ACCAATATGC CTGACGAAGG 480
CGTCATGTGC ACCAAATTCT ACTTCCAGGA GGACGAGTAG TACTGCCAGG GCCTGCCTGT 540
TCCCATTTCT GCATGGCAAG GACTGCAGGG ACTGCCAGTC CCCTGCCCCC AGGCTCCCG 600

```

	GCTATGGGG	CACCTGAGGAC	CAGCCATTGA	GGGGTGGACC	CTCAGAAGGC	GTCAACAACA	660
	CCTGGTCACA	GGACTCTGCC	TCCTCTTCAA	CTGACCAGCC	TCCATGCTGC	CTCCAGAATG	720
	GTCTTTCTAA	TGTGTGAATC	AGAGCACAGC	AGCCCTTGCA	CAAAGCCCTT	CCATGTCGCC	780
5	TCGTCAATCA	GGATCAAAAC	CCGACCACCT	GCCCAACCTG	CTCTCTCTT	GCCACTGCCT	840
	CTTCTCCCT	CATTCCACCT	TCCATGCCCC	TGGATCCATC	AGGCCACTTG	ATGACCCCA	900
	ACCAAGTGGC	TCCACACCC	TGTTTACAA	AAAAGAAAAG	ACCAGTCCAT	GAGGGAGGTT	960
	TTTAAGGGTT	TGTGAAAAT	GAAAATTAGG	ATTTCATGAT	TTTTTTTTT	CAGTCCCGT	1020
	GAAGGAGAGC	CCTTCATTG	GAGATTATGT	TCCTTCGGGG	AGAGGCTGAG	GACTTAAAT	1080
10	ATTCTGCAT	TTGTGAAATG	ATGGTGAAAG	TAAAGTGTAG	CTTTCCCTT	CTTTTCTTC	1140
	TTTTTTTGTG	ATGTCCCAAC	TTGTAAAAAT	TAAAGTTAT	GGTACTATGT	TAGCCCATATA	1200
	ATTTTTTTTT	TCCTTTTAAA	ACACTTCCAT	AATCTGGACT	CCTCTGTCCA	GGCACTGCTG	1260
	CCCAGCCTCC	AGCTCCATC	TCCACTCCAG	ATTTTTTACA	GCTGCTGCA	GTACTTTACC	1320
	TCCTATCAGA	AGTTTCTCAG	CTCCCAAGGC	TCTGAGCAAA	TGTGGCTCCT	GGGGTTCCT	1380
	TCCTCTCTG	CTGAGGAAT	AAATTGCTCC	TTGACATTGT	AGAGCTCTG	GCACTTGAG	1440
15	ACTTGTATGA	AAGATGGCTG	TGCTCTGCCC	TGTCTCCCCC	ACCAGGCTGG	GAGCTCTGCA	1500
	GAGCAGGAAA	CATGACTCGT	ATATGTCTCA	GGTCCCTGCA	GGGCCAAGCA	CCTAGCCTCG	1560
	CTCTTGGCAG	TACTCAGCG	AATGAATGCT	GTATATGTTG	GGTGCAAGT	TCCTACTTC	1620
	CTGTGACTTC	AGCTCTGTT	TACAATAAAA	TCTTGAATAA	GCCTAAAAAA	AAAAAAAATA	1680
20	AAAAAAAATA	AAAAAAAATA	AAAAAAAATA	AAAAAA			

Seq ID NO: 493 Protein sequence

Protein Accession #: NP_000568.1

25	1	11	21	31	41	51	
	MALETICRPS	GRKSSKMQAF	RINDVNQKTF	YLRNNQLVAG	YLQGPVNLE	EKIDVVPPIEP	60
	HALFLGHGG	KMCLSCVKSG	DETRLQLEAV	NITDLSENK	QDKRFAPIRS	DSGPTTSFES	120
	AACPFWLCT	AMEADQVSL	TNMPDEGVMV	TKFYQKDE			

Seq ID NO: 494 DNA sequence

Nucleic Acid Accession #: NM_002081.1

Coding sequence: 222..1898

35	1	11	21	31	41	51	
	GGCTGCCCGA	GOGAGCGTTC	GGACCTCGCA	CCCCGCGCGC	CCGCGCGCGC	CGCGCGCGCC	60
	GGCTTTTGT	GTCTCCGCT	CCTCGGCGCG	CGCGCGCTCT	GGACGCGGAG	CGCGCGCGCG	120
	CGGAGCCTTG	GCTCTGCGCT	TGCGGGGCGG	GAACTGCGCA	GGACCGCGCC	AGGATCGGAG	180
	AGAGGCGCGG	GCGGGTGGCC	GGGGGCGCGG	CCGGCCCGCG	CATGGAGCTC	CGGGCCCGAG	240
40	GCTGGTGGCT	GCTATGTGGG	GCCGAGCGCG	TGGTGGCGTG	CGCGCGCGGG	GACCGCGCCA	300
	GCAAGAGCGG	GAGCTGCGCG	GAGGTGCGCC	AGATCTAAGG	AGCCAAGGGC	TTGAGCCTGA	360
	GGGAGCTGGG	CCAGGCGGAG	ATCTGGGGTG	AGCACTCGCG	GATCTGTCCC	CAGGGCTACA	420
	CCTGCTGCAC	CAGCGAGATG	GAGGAGAACC	TGGCCAACCG	CAGCCATGCC	GAGCTGAGAA	480
	CGCGCTCCCG	GGACAGCAGC	CGCGTCTGCG	AGGCCATGCT	TGCCACCCAG	CTGCGCAGCT	540
45	TGATGACCA	CTTCAGACAC	CTGCTGAACG	ACTCGGAGCG	GACGCTGAGG	GCCACTTCC	600
	CGCGCGCTCT	CGGAGAGCTG	TACACGACAG	ACGCGAGGGC	CTTCCGGGAC	CTGACTCAG	660
	AGCTGCGGCT	GTACTACCGG	GGTGCCAACC	TGCACTGGA	GGAGACGCTG	GCGGAGTTCT	720
	GGGCGCGGCT	GCTCGAGCGC	CTCTTCAAGC	AGCTGCACCC	CCAGCTGCTG	CTGCTGATG	780
	ACTACTTGA	CTGCTCGGCG	AAGCAGGCGG	AGGCGCTGCG	GCCCTTCCGG	GAGGCCCGCA	840
50	GAGAGCTGCG	CCTGCGGGCC	ACCGTGGCTT	TGCTGGCTGC	TGCTCTCTTT	GTGAGGGGCC	900
	TGGGCGTGGC	CAGCGAGCTG	GTCCGGAAAG	TGGCTCAGGT	CCCGCTGGGC	CGGAGTGTCT	960
	CGAGAGCTGT	CATGAAGCTG	GTCTACTGTG	CTCACTGCTT	GGGAGTCCCG	GGCGCCAGGC	1020
	CCTGCCCTGA	CTATTGCCGA	AATGTGCTCA	AGGGCTGCCT	TGCCAACCAG	GCCGACCTGG	1080
	ACGCCGAGTG	GAGGAACCTC	CTGGACTCCA	TGTTGCTCAT	CACCGACAAG	TTCTGGGGTA	1140
55	CATCGGCTGT	GGAGAGTGTG	ATCGGCAGCG	TGCACACGTG	GCTGGCGGAG	GCCATCAAGC	1200
	CCCTCCAGGA	CAACAGGAGC	ACGCTCAAGG	CCAAGTCTAT	CCAGGCTGCG	GGGAACCCCA	1260
	AGGTCAACCC	CAGGCGGCTT	GGGCTGAGG	AGAAGCGCGG	CGGGGCAAGC	CTGGCCCGCG	1320
	GGGAGAGGCC	ACCTTCAGGC	ACGCTGGAGA	AGCTGGTCTC	TGAAGCCAAG	GCCAGCTCC	1380
	CGGAGCTCCA	GGACTTCTGG	ATCAGCCTCC	CAGGGACACT	GTGAGTGAAG	AAGATGGCCC	1440
60	TGAGCACTGC	CAGTATGAC	CGCTGCTGGA	ACGGGATGGC	CAGAGGCGCG	TACTCCCGCG	1500
	AGGTATGTTG	TGAGCGGCTG	GCCAAACCAG	TCAACAACCC	CGAGGTGGAG	GTGGACATCA	1560
	CCAAGCCGGA	CATGAACCTC	CGCAGCAGA	TCACTGACCT	GAAGATCATG	ACCAACCGCG	1620
	TGCGCAGCGC	CTACAACGCG	AACGACGTGG	ACTTCCAGGA	CGCCAGTGAC	GACCGCAGCG	1680
	GCTCGGCGAG	CGGTATGCGC	TGCTTGGATG	ACCTCTGGCG	CGGGAAGGTC	AGCAGGAAGA	1740
65	GCTCCAGTGC	CGGAGCGGCC	TTGACCCATG	CCCTCCAGAG	CCTGTCAGAG	CAGGAAGGAC	1800
	AGAAGACCTC	GGCTGCCAGC	TGCCCCAGCG	CCCGAGCTTT	CCTCTGCCCC	CTCTCTCTCT	1860
	TCCTGGCCCT	TACAGTAGCC	AGGCCCCCGT	GGCGGTAACT	GCCCCAAGGC	CCGAGGGACA	1920
	GAGGCCAAGG	ACTGACTTTG	CCAAAAATAC	AACACAGACG	ATATTTAATT	CACCTCAGCC	1980
	TGGAGAGGCC	TGGGGTGGGA	CAGGGAGGGC	CGCGCGCTCT	GAGCAGGGGC	AGGCGCAGAG	2040
70	GTCCGAGGCC	CAGGCTGGGC	CTCGCCTGCC	TTTCTGCTTT	TTAATTTTGT	ATGAGTCTCT	2100
	CAGGTCAAGT	GGGAGCAGT	GTGCCCAAAA	GCCATGTATT	TCAGGGAAGT	CAGGGGCAAC	2160
	TCGGGCTGCC	TAGCCCTCCC	CCCAGCTCCC	TGCACCGCGG	CAGAAGCAGC	CCCTCAGGCG	2220
	CTACAGAGGA	GGCTCTAAAG	CAACCCGCTG	GAGCCACAGG	CGAGCCTGTG	CCTTCTCTCC	2280
	CGCTCTCTCC	CAGTGGGACT	CCCAGCAGAG	CCCACACAGC	AGCCCTGGCC	CACCCCCAG	2340
75	CCTCCAGAGA	AGCCCCGAC	GGGCTGTCTG	GGTGTCCGCC	ATCCAGGGTC	TGGCAGAGCC	2400
	TCTGAGATGA	TGATGATGCG	CCTCCCTTCA	GCGCAGGCTG	CAGAGCCCGG	CCCCACCTCC	2460
	CTGCGCCCTT	GAGGGGCCCC	AGCGTCTGCA	GGGTGACGCC	TGAGACAGCA	CCACTGCTGA	2520
	GGAGTCTGAG	GAGTGTCTCT	CCACAGACCC	TGCAGTGAAG	GGCCCTCCAT	GCGCAGATGA	2580
	GGGGCCACTG	ACCCACCTGC	GCTTCTGCTG	GAGGAGGGGA	AGCTGGGCCC	AAAGGCCCAG	2640
80	GAGGCGAGCG	TGGGCTCTGC	CAATGTGGGC	TGCCCTCCGC	ACACAGGGCT	CACAGGGCAG	2700
	GCCTTGTCTG	GCTCCAGGCG	TGTTGGAGGA	CCCGAGGGGC	TGAGGAGCAG	CCAGGACCCG	2760
	CCTGCTCCCA	TCTTCAACCA	GATCAGGAAC	CAGGGCCTCC	CTGTTCAAGG	TGACACAGGT	2820
	CAGGGCTCAG	GAGTCAACCT	GGCTGTACCC	TGCTCACAGG	GATGCTGGTG	GCTGGTGAGA	2880
	CCCGCAGCTG	CACACGGGAA	TGCCATAGTC	CCTTCCCGAC	CAGGCCAGCT	GCACTGCAAG	2940
85	GCAAGGGGAC	CTGGATAGTT	AAGGGCTTTT	CCAAACATGC	ATCCATTATC	TGCACTTCC	3000
	TGTCTTGTGT	CATGAGAGGC	TGTTGCTCTC	TCCAGATGGG	CTTCGAGGCG	CGCAGGGGCC	3060
	CACCTTGGAC	CCTGTGACCT	TCTGTCACT	CACTGAGGCC	ATCAGGGCCC	TGCCCCAGGC	3120

5 CTGGACGGGC CCTCCTTCCC TCCTGTGCCC CAGCTGCCAG GTGGCCCTGG GGAGGGGTGG 3180
TGTGGTGTGG GGAAGGGGTC CTGCAGGGGG AGGAGGACTT GGAGGGTCTG GGGGCAGCTG 3240
TCCTGAACCG ACTGACCCTG AGGAGGGCGC TTAGTGCTGC TTTGCTTTTC ATCACCGTCC 3300
CGCACAGTGG ACGGAGGTCC CCGGTGCTG GTCAGGTCCC CATGGCTTGT TCTCTGGAAC 3360
CTGACTTTAG ATGTTTTTGG ATCAGGAGCC CCCAACACAG GCAAGTCCAC CCCATAATAA 3420
CCTGCGCACT GCCAGGGTGG GCTGGGGACT CTGGCACAGT GATGCCGGGC GCCAGGACAG 3480
CAGCACTCCC GCTGCACACA GACGGCCTAG GGGTGGCGCT CAGACCCAC CCTACGCTCA 3540
TCTCTGGAAG GGGCAGCCCT GAGTGGTCAC TGGTCAGGGC AGTGGCCAAG CCTGCTGTGT 3600
10 CCTTCCTCCA CAAGGTCCCC CCACCGCTCA GTGTCAGCGG GTGACGTGTG TTCTTTTGAG 3660
TCCTTGATG AATAAAGGC TGGAAACCTA AA

Seq ID NO: 495 Protein sequence
Protein Accession #: NP_002072.1

15 1 11 21 31 41 51
MELRARGWNL LCAAAALVAC ARGDPASKSR SCGEVRQIYG ARGFSLSDVP QAEISGEHLR 60
ICPGYTCCT SEMEENLANR SHALETALR DSSRVLQAML ATQLRSFDDH PQHLNDSESR 120
20 TLQATFFGAF GELYTONARA FRDLYSELRL YRGNLHLE ETLAEFWARL LERLFKQLHP 180
QLLLPDDYLD CLGQAEALR PFGAPRELRL LRATRAFVA RSVFQGLGVA SDVVRKVAQV 240
PLGPECSRVA MKLYCAHCL GVPGARPCPD YCRNVLKGL ANQADLDAEW RNLLDSMWLI 300
TDKFWGTSGV ESVGSGVHTW LAEAINALQD NRDLTAKVI QGCGNPKVNP QGPGPEEKRR 360
RGKLAPRERF PSGTLEKLVS EAKAQLRDVQ DFWISLPGLT CSEKMALSTA SDDRWNWMA 420
25 RGRYLPVEMG DGLANQINN P EVEVDITKPD MTRIQIMQL KIMTNRLRSA YNGNDVDFQD 480
ASDDGSGSGS

Seq ID NO: 496 DNA sequence
Nucleic Acid Accession #: NM_001650.2
Coding sequence: 40..1011

30 1 11 21 31 41 51
GGGGCAGGCA ATGAGAGCTG CACTCTGGCT GGGGAAGGCA TGAGTGACAG ACCCACAGCA 60
AGGGGTGGGG GTAAGTGTGG ACCTTTGTGT ACCAGAGAGA ACATCATGGT GGCCTTCAAA 120
35 GGGGTCTGGA CTCAAGCTTT CTGGAAGCA GTCACAGCGG AATTCTGGGC CATGCTTATT 180
TTTGTCTCC TCAGCCTGGG ATCCACCATC AACTGGGGTG GAACAGAAAA GCCTTTACCG 240
GTGACATGG TTCTCATCTC CCTTTGCTTT GGACTCAGCA TTGCAACCAT GGTGCAGTGC 300
TTTGGCCATA TCAGCGGTGG CCACATCAAC CCTGCAGTGA CTGTGGCCAT GGTGTGCACC 360
AGGAAGATCA GCATCGCCAA GTCTGTCTTC TACATCGCAG CCCAGTGCCT GGGGGCCATC 420
40 ATTGGAGCAG GAATCCTCTA TCTGTCCACA CCTCCAGTG TGGTGGGAGG CTGGGAGTGC 480
ACCATGTTTC ATGGAATCTT TACCGCTGGT CATGGTCTCC TGGTTGAGTT GATAATCACA 540
TTTCAATTGG TGTTTACTAT CTTTCCAGC TGTGATTCCA AACCGACTGA TGTCACTGGC 600
TCAATAGCTT TAGCAATTGG ATTTTCTGTT GCAATTGGAC ATTTATTGTC AATCAATTAT 660
ACTGGTGCCA GCATGAATCC CGCCGATCC TTTGGACCTG CAGTTATCAT GGGAAATTGG 720
45 GAAACCAATT GGATATATTG GGTGGGGCCC ATCATAGGAG CTGTCTCTCG TGGTGGCCTT 780
TATGATATG TCTTCTGTCC AGATGTTGAA TTCAAACGTC GTTTTAAAGA AGCCTTCAGC 840
AAAGCTGCCC AGCAAAACAA AGGAAGCTAC ATGGAGGTGG AGGACAACAG GAGTCAGGTA 900
GAGACGGATG ACCTGATTCT AAAACCTGGA GTGGTGCAATG TGAATTGACGT TGACCGGGGA 960
GAGGAGAAGA AGGGGAAAGA CCAATCTGGA GAGGTATTGT CTTCAATATG ACTAGAAGAT 1020
50 GCACTGAAA GCAGACAAGA CTCCTTAGAA CTGTCTCAG ATTTCTCTCC ACCCATTAAG 1080
GAAACAGATT TGTATATAAT TAGAAATGTG CAGGTTTGT GTTTCATGTC ATATTACTCA 1140
GTCTAATGAA TAAATATTTC ATAATTIACA AAGGAGGAAC GGAAGAAGCC TATTGTGAAT 1200
TCCAAATCTA AAAAAGAAA TATTTTAAAG ATGTTCTTAA GCAATATAT ACCTATTTTA 1260
TCTAGTTACC TTTCAATTAAC AACCATTITT AACCGTGTGT CAAGATTGG TTAAGTCTTG 1320
55 CCTGACAGAA CTCAAAGACA CGTCTATCAG CTATTCTCT CTCTACTGGA ATATTGGTAT 1380
AGTCAATTCT TATTGAATA TTTATTCTAT TAAACTGAGT TTAACAATGG C

Seq ID NO: 497 Protein sequence
Protein Accession #: NP_001641.1

60 1 11 21 31 41 51
MSDRPTARRW GKCGPLCTRE NIMVAFKGVN TQAFWKAVTA EFLAMLIFVL LSLGSTINWG 60
GTEKPLVDVM VLISLCPGLS IATMVQCFGH ISGGINPAV TVAMVCTRKI SIAKSVFYIA 120
65 AQCLGAIIGA GILYLVTPPS VVGLGVIMV HGNLTAGHGL LVELIITFQL VFTIFASCDS 180
KRTDVTGSIA LAIGFSVAIG HLFALNTGA SMNPAPSPGP AVIMGNWENH WIYVWGPIIG 240
AVLAGGLYFV VFCPDVEFKR RFKEAFSKAA QQTGKSYMEV EDNRSGVETD DLILKPGVVH 300
VIDVDRGEER KKKQSGEVL SSV

Seq ID NO: 498 DNA sequence
Nucleic Acid Accession #: AB020684.1
Coding sequence: 1..1744

75 1 11 21 31 41 51
CCCCCTTGT ATTAAATACAT TAAAAAGATT CAATCTTTAC CCTGAGGTAA TTTTGGCCAG 60
TTGTATACCG ATTATACCA AAATAATGGA CTTGATTGGT ATTCAAACCA AGATATGTTG 120
GACGGTTACC AGAGGAGAAG GACTCAGTCC TATTGAAAGC TGTGAAGGAT TGGGAGATCC 180
TGCTTGCTTT TATGTTGCTG TAAATTTTAT TTTAAATGGA CTAATGATGG CATTTATTCTT 240
80 CATATATGGC GATATATTAA GTGGCAGCCG ATTAGGAGGC CTGGTTACAG TGTGTGCTTT 300
CTTTTCAAT CATGGAGAGT GTACCCGTGT AATGTGGACA CCACCTCTCC GTGAAAGCTT 360
CTCATATCCA TTTCTTGTTC TTCAGATGTT GCTAGTGACT CATATTCTCA GGGCTACAAA 420
ACTTTATAGA GGAAGCTTGA TTGCACTCTG CATTTCOAAT GTATTTTTC TGTCTCTCTG 480
GCAGTTTGCT CAGTTTGTAC TTCTTACTCA GATTGCATCA TTATTTCAG TATATGTTGT 540
85 CGGATACATT GATATATGTA AATTACGGAA GATCATTAT ATACACATGA TTTCTCTTGC 600
ACTTTGTTTT GTTTGTAGT TTGGGAACCTC AATGTTATTA ACTTCTTATT ATGCTTCTTC 660
TTTGTAATT ATTTGGGGTA TTCTGGCAAT GAAACACAT TTCTGAAAA TAAATGTATC 720

5
10
15
20
25
30
35
40
45
50
55
60

TGAACCTAGT TTATGGGTGA TTCAAGGATG TTTTGGTTA TTTGGAACTG TCATACCTAA 780
 ATACTTGACA TCTAAAATTT TTGGTATTGC AGATGACGCT CATATTGGCA ACTTACTAAC 840
 ATCAAAATTC TTTAGTTATA AGGATTTTGA TACTTTATTG TATACCTGTG CAGCGGAGTT 900
 TGACTTTATG GAAAAGAGAGA CTCCACTGAG ATACACAAAG ACATTATTGC TTCCAGTTGT 960
 TCTTGTAGTG TTTGTGCTA TTGTTAGAAA GATTATTAGT GATATGTGGG GTGTCTTAGC 1020
 TAAACACAG ACACATGTAA GAAAACACCA GTTTGATCAT GGAGAGCTGG TTTACCATGC 1080
 ATTGCAATTG TTAGCATATA CAGCCCTTGG TATTTTAATT ATGAGACTAA AACTCTTCTT 1140
 GACACACAC ATGTGTGTGA TGGCATCACT GATCTGCTCA AGACAGCTAT TTGGATGGCT 1200
 CTTTGTGCAA GTACATCTTG GTGCTATTGT GTTGTCTATA TTAGCAGCAA TGTCAATACA 1260
 AGGTTTCAGCA AATCTGCAA CCCAGTGGAA TATTGTAGGG GAGTTCAGCA ATTTGCCCCA 1320
 AGAAGAACTT ATAGAATGGA TCAATATAG TACTAAACCA GATGCAGTGT TTGCGGGTGC 1380
 CATGCCACAG ATGGCAAGTG TTAAGCTCTC TGCACCTCGG CCCATTGTGA ATCATCCACA 1440
 TTATGAAGAC GCAGGCTTGA GAGCCAGAAC AAAAATAGTA TACTCAATGT ATAGTCGGAA 1500
 AGCAGCCGAA GAAGTGAAGC GAGAACTGAT AAAGTTAAAA GTGAATATT ACATTCTAGA 1560
 AGAGTCATGG TGTGTAAGAA GATCCAAGCC TGGTTGCACT ATGCCTGAAA TTTGGGATGT 1620
 AGAAGATCCT GCCAATGCTG GAAAACTCC CTTATGTAAC CTCTTGGTGA AGGATTCCAA 1680
 ACCTCACTTC ACCACTGTAT TCCAGAACAG TGTTTACAAA GTCCTAGAAG TTGTAAGAAG 1740
 ATGACTGCTA CATGACCTGC TGCCTACGGA GAATACATC TGTATGGTT TTAATGTTTT 1800
 GCTAAGTCAT GTGTGTGTCA TATCCCAAAA ACTTTTATAG GTAAGTGT TTCAATAGAA 1860
 AAGCTTTTAT TTGCTCAATT TGAATGTAT TCTAATTATA AAAATGACTT ACACCTTTAT 1920
 CAATTGGTGA CTATTTCAAT GCACCTTTTA AAAATTGCTA TGCAATGAG TATATGCTTG 1980
 TACTTGACTT TAATATTGTG GCTAAAGTGA GCAAGCTTAC CTGTATAAAG AAAACACAGT 2040
 GGGTTGTGAC AAGGATGACA TGAATAACA GGCAATTTCT GACATGTAG GGGCTGATTT 2100
 TATAGTGTA GAACATTATA TGCCCTTTC TTTCTTTTTC TGCCCTCTGC TCTTGTCTTT 2160
 TGGACATTTT TGGTATGTA AGTTCTTCGG TCAATGTCAGC COCTGTCTAT AACTTGAGTT 2220
 ACAGTAGATG GGGCAGACAT GGAGTGTGTT CTATATAAAA CTATCTGTTT GTTTTACTTC 2280
 CTGTGTGCGT TTTGTCTCTC TGTCTCTTGT TTAATGAAGC TTTTCTGCC CATTATTAAT 2340
 CCAAACTCTT GGACCTTGTG GTTAGGAAAT TCCTTAACT TCCAGCCATA TGGCATTATC 2400
 GTGTCTCTTT CTCTCTCTCT CTGTCTCTCT CTCTCTCTCT CTTCGCCATA TTTTCTGTCA 2460
 AATAAGTACT GTTTACTCAT TTAGTGTCTT ATCAAGTACT TATCTTGGT TTTAAAAAAA 2520
 ATTAATGGTA ACTGATTTT TCTCATTTT AGCATTATTC AAATGTTTAT ATTTTAAATAC 2580
 CTTTAAACCA CTCTAAAGT TTTTCATGTT TAATTATAGT TTTAAGAAA ACTATTTTGA 2640
 ACAACCCCAA ATATAGTGCA TCTAGAACT AATGTATATT TGATTAGACA TCATTTTATAG 2700
 TGAACAGTA GACTGTAGTA CATGGTAATT TTTCTTTTAC TATTAAGATA CAATAAAACA 2760
 TGACTAATT TTGCTCAAA AATGTAAAGA ATAATGATA ATGGAGTTT TATATTTTA 2820
 CTTTAAAGAT TGCGTGTCT TAATAAGACA AAGCCTTAAG CCTTATGTTA TAATTTGGT 2880
 TCTAAAAACC ATCATTTAG TATAAGGAAT AAGTATATT CGTCTCTCTC TTTAGTTTTT 2940
 TTCTTCTAT TATTTTTTAT TTTGAAAAAT TTCTACACCT TCTTTGAATT CCTTGTATGA 3000
 ATTTTTGTTT CTTAGAGATT AATTGTGTG AATGAGATT CTTCAAAACG ATGAAACCTC 3060
 ATAGCTCTGA GAAAAGGTT TAGGGTTTTA AATCTAAGC AAAGCGTGAC TATGGCTGAC 3120
 AGACTACACA TTAATTTATA CAGCTTCTCT TTTCTAACA CAGGCAGATT AACCTCATG 3180
 TGGATTGTCC TTCAGACCTT AGTCTCAGC CATGTTTCT GGTGCCCATC CTGGAAGCC 3240
 GCTGTTCCTT TTCTACCTTC TTACCAGAGC CCAAGGGCAG GCCTGGTCCC GGGGAAGCAG 3300
 CAGCTTGCTG ACATAAGTCA GCTGCAAGG CTGAGGAGTG TGCCCTCAGA GAAGCACCGC 3360
 CCCCCAGTCT TGTGCTCAGC CCTAGAGCCG CAGCTCCAG GGATGCTCCT TCCCTGGAGG 3420
 CAGCCACAGA GAGGAGCTCT GGCAGCGTTC TTGAGATTG TGGCCACTGT TTCTCATTTG 3480
 CTGTTGACT GTTTTATT TTTAGGCTTT TGCTAGTTT AGAAAATAGG GAAGCAGCCC 3540
 TTGATTGTG GATTAAAGC AACATTGAG CGATGATGCA CAACAGTCCA GGAATATGGG 3600
 CGGTGGACAC TTGAGGCTGA GGATGGGAGT TGACATGAGC AGGGAGAGGG AGGTGCGCGC 3660
 TGCTTATCTG TGAATTGTG TCACCTGAGT GTGGCTGATT GTGTACATCC AGCAGTTACA 3720
 ATTTTAAAA ATTATCTTT TACATTTATT TTTATTTTT CTACCCCCA GTAAITTCCT 3780
 TCCAAAGAG TTACATGTA ATAAGTAGAA ATTCTGTATA GGA AAAAGC ATTTAAAAA 3840
 CTATTATAAC TGCTTCATT GCTGGGAACC ATTAAGATA ATATAAATA GCTTTTCCA 3900
 GAAGGATCCT TTTGTAGCAG TGTTTATGAA TGTAACCCCC AGCAAAATAT GGCTATATAT 3960
 TAGGGAGAGC AGTTTGGAGC AGAGGCTGTA AGGTCCCTGC TATGCAGCCG TGGCCACAGC 4020
 TCGCAGCCCA AGCATGTGG AGCATCCACA CCTTGTATG CAATGCAGAT TGGTAGCAGG 4080
 TTCCATAGGC GTACAAAAC GTATTAAAG TCAGTGTGTT GCATATTGTT AGCATTTACA 4140
 AATATTTTTG CTTTAGTAT AGGAAAGTAA GGATGGGCAA AGAAGCGATC AAAATAGCTA 4200
 TTGCTACAC ATTTTCAAAA ACAAGTTGG GGCCTGATT CTTTAAAAAG ATAAGCCTCT 4260
 AAAAAAGCTT GGGCAAAAAA ATATAGTGT AAAATAGGCC AGTGATATTA ATGAGAAAA 4320
 GAAAGTATGT ATCAGGAATA AAGTATATT GCATAGGAGT ATTGTATTTT TATGAATTTT 4380
 ATGCCAGTTG TTTACATGTA CTATATATGT TAAATTA AAAATCATGA GAAATG

65 Seq ID NO: 499 Protein sequence
 Protein Accession #: BAA74900.1

70
75
80

1 11 21 31 41 51
 PLVINTLKRF NLYPEVILAS WYRIYTKIMD LIGIQTKICW TVTRGEGLSF IESCEGLGDP 60
 ACFYVAVIFI LNLMMALFF IYGYLSGSR LGGLVTVLCF FPNHGECTRV MNTPLRESF 120
 SYPFLVLQML LVTHILRATK LYRGSIALC ISNVFFMLPW QFAQFVLLTQ IASLEAVYVV 180
 GYIDICKLRK IYIHMISLA LCFVLMFGNS MLLTSYASS LVIIWGLIAM KPHFLKINS 240
 ELSLWVIQCG YTKTGLVILK YLTSKIPGIA DDAHIGNLLT SKFPSYKDFD TLLYTCAAEF 300
 DFMEKETPLR YTKTLLPVV LVVFAIVRK IISDMWGLA KQOTVVRKHQ FDHGLVYHA 360
 LQLLAYTALG ILMRLKFL TPBMCMVASL ICSRLFGWL PCKVHPGAIV FAILAAMSIO 420
 GSANLQTNW IVGEFNSLPQ EELIEWIKYS TKPDAVFAGA MPTMASVKLS ALRPIVNHPH 480
 YEDAGLRART KIVYSYMRK AAEVVKRELI KLVKNYILE ESWCVRRSKP GCSMPFIWDV 540
 EDPANAGKTP LCNLLVKDSK PHTTVFQNS VYKVLVVKE

80 Seq ID NO: 500 DNA sequence
 Nucleic Acid Accession #: NM_001276.1
 Coding sequence: 127..1278

85
 1 11 21 31 41 51
 AGTGGAGTGG GACAGGTATA TAAAGGAAGT ACAGGGCCCTG GGAAGAGGCC CCTGTCTAGG 60
 TAGCTGGCAC CAGGAGCCGT GGGCAAGGGA AGAGGCCACA CCTGCGCTG CTCTGTCTGCA 120

5
10
15
20
25
30

```

GCCAGAATGG GTGTGAAGGC GTCTCAAACA GGCTTTGTGG TCCTGGTGCT GCTCCAGTGC 180
TGCTCTGCAT ACAAACTGGT CTGCTACTAC ACCAGCTGGT CCCAGTACCG GGAAGGCGAT 240
GGGAGCTGCT TCCAGATGCG CCTTGACCGC TTCTCTGTGA CCCACATCAT CTACAGCTTT 300
GCCAATATAA GCAACGATCA CATCGACACC TGGGAGTGGG ATGATGTGAC GCTCTACGGC 360
ATGCTCAACA CACTCAAGAA CAGGAACCCC AACCTGAAGA CTCTCTTGTG TGTCCGAGGA 420
TGGAACTTTG GGTCTCAAAG ATTTTCCAAG ATAGCCTCCA ACACCCAGAG TCGCCGGACT 480
TTCAATCAAG CAGTACCGCC ATTCTGCGC ACCCATGGCT TTGATGGGCT GGACCTTGCC 540
TGGCTCTACC CTGGACGGAG AGACAAACAG CATTTTACCA CCTAATCAA GGAAATGAAG 600
GCCGAATTTA TAAAGGAAGC CCAGCCAGGG AAAAAGCAGC TCCTGCTCAG CGCAGCACTG 660
TCTGCGGGGA AGGTCAACAT TGACAGCAGC TATGACATTG CCAAGATATC CCAACACCTG 720
GATTTCATTA GCATCATGAC CTACGATTTT CATGGAGCCT GGCCTGGGAC CACAGGCCAT 780
CACAGTCCCC TGTTCGAGG TCAGGAGGAT GCAAGTCTCG ACAGATTGAG CAACACTGAC 840
TATGCTGTGG GGTACATGTT GAGGCTGGGG GCTCCTGCCA GTAAGCTGGT GATGGGCATC 900
CCCACCTTGG GAGCAGGCTT CACTCTGGCT TCTTCTGAGA CTGGTGTGGG AGCCCCAATC 960
TCAGGACCGG GAATTCACCG CCGGTTTACC AAGGAGGCAG GGACCCCTGC CTACTATGAG 1020
ATCTGTGACT TCCTCCGCGG AGCCACAGTC CATAGAACCC TCGGCCAGCA GGTCCCCTAT 1080
GCCACCAAGT GCAACCAAGT GGTAGGATAC GACGACCAAG AAAGCGTCAA AAGCAAGGTG 1140
CAGTACCTGA AGGATAGGCA GCTGGCAGGC GCCATGGTAT GGGCCCTGGA CTGTGATGAC 1200
TTCACGGGCT CCTTCTGCGG CCAGGATCTG CGCTTCCTTC TCACCAATGC CATCAAGGAT 1260
GCACCTGGCT CAACGTAGCC CTCTGTTCTG CACACAGCAC GGGGGCCAAG GATGCCCGGT 1320
CCCCCTCTGG CTCCAGCTGG CCGGAGCCTT GATCACTGCG CCTGCTGAGT CCCAGGCTGA 1380
GCCTCAGTCT CCTCCCTTGG GGGCCTATGC AGAGGTCCAC AACACACAGA TTTGAGCTCA 1440
GCCCTGGTGG GCAGAGAGGT AGGGATGGGG CTGTGGGGAT AGTGAGGCAT CGCAATGTAA 1500
GACTCGGGAT TAGTACACAC TTGTTGATGA TTAATGGAAA TGTTTACAGA TCCCCAAGCC 1560
TGGCAAGGGA ATTTCTTCAA CTCCTTGCCC CTATACAAAG GACACCATTT 1620
TGGCAAGCTG TATCACCAGG GAGCCAAACA TCCTACAGA CACAGTGACC ATACTAATTA 1680
TACCCCTGCG AAAGCCAGTG TGAAAACCTTC ACTTAGGAAC GTAATCGTGT CCCCTATCCT 1740
ACTTCCCTTT CCTAATTCOA CAGCTGCTCA ATAAAGTACA AGAGTTTAAC AGTGTGTTGG 1800
CGCTTTGCTT TGGTCTATCT TTGAGCGCCC ACTAGACCCA CTGGACTCAC CTCCCCCATC 1860
TCTTCTGGGT TCCTTCTCTT GAGCCTTGGG ACCCTGAGC TTGCAGAGAT GAAGGCCGCC 1920
ATGTT

```

Seq ID NO: 501 Protein sequence
Protein Accession #: NP_001267.1

35
40
45

```

1 11 21 31 41 51
MGVKASQTGF VVLVLLQCCS AYKLVCYYTS WSQYREGDGS CFPDALDRFL CTHIIYSPAN 60
ISNDHIDTWE WNDVTLYGML NTLKRNPNL KTLISVGGWN FGSQRFPSKIA SNTQSRRTFI 120
KSVPPFLRTH GFDGLDLAWL YPGRRDRQHF TTLIKEMKAE PIKEAQPGKK QLLLSAALSA 180
GKVTIDSSYD IAKISQHLDF ISIMTYDFHG AWRGTTGHHS PLFRGQEDAS PDRFSNTDYA 240
VGVMRLRGAP ASKILVMGIDP FGRSFTLASS ETGVGAPISG PGIPGRFTKE AGTLAYYEIC 300
DFLRGATVHR TLGQQVPIAT KGNQWVGYYD QESVSKVQY LKDRQLAGAM VWALDLDDFQ 360
GSFCGQDLRF PLTNALKDAL AAT

```

Seq ID NO: 502 DNA sequence
Nucleic Acid Accession #: NM_006474.1
Coding sequence: 181..669

50
55
60
65

```

1 11 21 31 41 51
GCTGCCTAGG GTCTGAAAG CTGGGGCACC CTCCTCTCC GGGGCTCCTG CTCOCACCCC 60
TCCGGCCCCC CCACCGTCGC GCTCCTCCAG GCTGGGCTCG TGGCCGGCGT GCTTTTAATT 120
TTCCCCCAGC TCAGAACTCT GCTGCTCGGC CCCAGGAGA GCAACAACTC AACGGGAAG 180
ATGTGGAAGG TGTCACTCT GCTCTTCGTT TTGGGAAGCG CGTGGCTCTG GGTCTGGGCA 240
GAAGGAGCCA GCACAGSCCA GCCAGAAGAT GACACTGAGA CTACAGGTTT GGAAGGCGGC 300
GTTGCCATGC CAGGTGCCGA AGATGATGTG GTGACTCCAG GAACCAAGCA AGACCGCTAT 360
AAGTCTGGCT TGCAACTCT GGTGGCAACA AGTGTCAACA GTGTAACAGG CATTCGCATC 420
GAGGATCTGC CAATCTCAGA AAGCAGTGC CACGCGCAAG AACAAAGTCC AAGCGCCACA 480
GCCTCAAGCG TGGCCACCA TCACTCCAG GAGAAAGTGG ATGGAGACAC ACAGACAACA 540
GTTGAGAAAG ATGTTTGTG AACAGTGACC CTGGTTGGAA TCATAGTTGG GGTCTTACTA 600
GCCATCGGTT TCATTGGTGG AATCATCGTT GTGGTTATGC GAAAAATGTC GGAAGGTIAC 660
TCGCCCTAAA GAGCTGAAGG GTTACGCCCT GCTTGCCAAC GTGCTTTAAA AAAAGACCGT 720
TTCGACTCT GTGGCCCTGT CCCTGAGCTC GTGGGGAGAA GATGACCCCT GGAACATTGT 780
CGGGCCCATC CAGATTCCAC GGTGACTTTC CGTTTGCCAA ATTAACCGAG GAAAGACCTT 840
TCACCAGATT TGGTCTTAA ACTTT

```

Seq ID NO: 503 Protein sequence
Protein Accession #: NP_006465.1

70
75

```

1 11 21 31 41 51
MWKVSALLFV LGSASLWVLA EGASTGQPED DTETTGLEGG VAMPGAEDDV VTPGTSEDRY 60
KSGLTTLVAT SVNSVTGIRI EDLPTSESTV HAQEQSPSAT ASNVATSEST EKVDGDTQTT 120
VERDGLSTVT LVGIIVGVLL AIGFIGGIIV VVMRKMSGRY SP

```

Seq ID NO: 504 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 62..895

80
85

```

1 11 21 31 41 51
CACTGCTCTG AGAATTGTG AGCAGCCCTT AACAGGCTGT TACTTCACTA CAACTGACGA 60
TATGATCATC TTAATTACT TATTTCTCTT GCTATGGGAA GACACTCAAG GATGGGGATT 120
CAAGGATGGA ATTTTTCATA ACTCCATATG GCTTGAACGA GCAGCCGGTG TGTACCACAG 180
AGAAGCACGG TCTGCAAAAT ACAAGCTCAC CTACCCAGAA GCTAAGGCGG TGTGTGAATT 240
TGAAGGCGGC CATCTGCGAA CTTACAAGCA GCTAGAGGCA GCCAGAAAAA TTGGATTCTA 300

```


5
10
15
20

```
TGCTCTGCT GCTGGATGGA TGGCTAAGGG CAGAGTTGGA TACCCCATTTG TGAAGCCAGG 360
GCCCAACTGT GCGATTTGGAA AAACTGGCAT TATTGATTAT GGAATCCGTC TCAATAGGAG 420
TGAAGATGGG GATGCTTATT GCTACAACCC ACAACGAAAG GAGTGTGGTG GCGTCTTTAC 480
AGATCCAAAG CAAATTTTTA AATCTCCAGG CTCCCAAAT GAGTACGAAG ATAACCAAAT 540
CTGCTACTGG CACATTAGAC TCAAGTATGG TCAGCGTATT CACCTGAGTT TTTTAGATTT 600
TGACCTTGAA GATGACCCAG GTTGCTTGGC TGATTATGTT GAAATATATG ACAGTTACGA 660
TGATGTCCAT GGCTTTGTGG GAAGATACTG TGGAGATGAG CTTCAGATG ACATCATCAG 720
TACAGGAAAT GTCATGACCT TGAAGTTTCT AAGTGATGCT TCAGTGACAG CTGGAGGTTT 780
CCAAATCAAA TATGTTGCAA TGGATCCTGT ATCCAAATCC AGTCAAGGAA AAAATACAAG 840
TACTACTTCT ACTGGAATA AAAACTTTTT AGCTGGAAGA TTTAGCCACT TATAAAAAAA 900
AAAAAAGGA TGATCAAAAC ACACAGTGTT TATGTTGGAA TCTTTTGGAA CTCCTTTGAT 960
CTCACTGTTA TTATTAACAT TTATTTATTA TTTTCTTAAA TGTGAAAGCA ATACATAATT 1020
TAGGGAAAT TGGAAATAT AGGAAACTTT AAACGAGAAA ATGAAACCTC TCATAATCCC 1080
ACTGCATAGA AATAACAAGC GTTAACATTT TCATATTTTT TTCTTTCACT CATTTTCTA 1140
TTTGTGGTAT ATGTATATAT GTACCTATAT GTATTTGCAT TTGAAATTTT GGAATCCTGC 1200
TCTATGTAGA GTTTTGTATT ATACTTTTAA AATCTTGAAC TTTATAAACA TTTTCTGAAA 1260
TCATTGATTA TTCTACAAA ACATGATTTT AAACAGCTGT AAAATATTCT ATGATATGAA 1320
TGTTTATGCT ATTATTAAG CCTGTCTCTA TTGTTGGAAT TTCAGGTCAT TTTCATAAAT 1380
ATTGTTGCAA TAAATATCCT TGAACACACA AAAAAAAAAA AA
```

Seq ID NO: 505 Protein sequence
Protein Accession #: Eos sequence

25
30

```
1 11 21 31 41 51
MIIILYLFLL LWEDTQGWGF KDGIFHNSIW LERAAGVYHR EARSQKYKLT YAEKAVCEP 60
EGGHLATYKQ LEAARKIGFH VCAAGWMAKG RVGYPIVKPG PNCXFGKTGI IDYGI RLNRS 120
ERWDAYCYNP HAKECCGVPT DPKRIPKSPG PFNEYEDNQI CYWHIRLKYG QRIHLSFLDF 180
DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELFPDDIIS TGNVMTLKFL SDASVTAGGF 240
QIKYVAMPDV SKSSQGNKNTS TTSTGNKNFL AGRPSHL
```

Seq ID NO: 506 DNA sequence
Nucleic Acid Accession #: NM_007115.1
Coding sequence: 69..902

35
40
45
50
55
60

```
1 11 21 31 41 51
GAATTGCGAC TGCTCTGAGA ATTTGTGAGC AGCCCTTAAC AGGCTGTGTTAC TTTACTACAA 60
CTGACGATAT GATCATCTTA ATTTACTTAT TTCTCTTGCT ATGGGAAGAC ACTCAAGGAT 120
GGGGATTCAA AGATGGGAAT TTTCATAACT CCATATGGCT TGAACGAGCA GCCGCTGTGT 180
ACCACAGAGA AGCAGCGTCT GGCAATACA AGCTCACCTA CGCAGAAGCT AAGGCGGTGT 240
GTGAATTTGA AGGCGGCAT CTGCAACTT ACAAGCAGCT AGAGGCAGCC AGAAAAATTG 300
GATTTTCATGT CTGTGCTGCT GGTGGATGG CTAAGGGCAG AGTTGGATAC CCCATTGTGA 360
AGCCAGGGCC CAACCTGATGA TTTGGAAAAA CTGGCATTAT TGATTATGGA ATCCGTCTCA 420
ATAGAGTGA AAGATGGGAT GCCTATTGCT ACAACCCACA CGCAAGGAGT TGTGTGGCG 480
TCCTTACAGA TCCAAAGCGA ATTTTAAAT CTCCAGGCTT CCCAAATGAG TACGAAGATA 540
ACCAATCTG CTACTGGCAC ATTAGACTCA AGTATGGTCA GCGTATTCCG CTGAGTTTGT 600
TAGATTTTGA CCTTGAAGAT GACCCAGGTT GCTTGGCTGA TTATGTTGAA ATATATGACA 660
GTTACGATGA TGTCCATGGC TTGTGTTGGA GATACCTTGG AGATGAGCTT CCAGATGACA 720
TCATCAGTAC AGGAATGTCT ATGACCTTGA AGTTTCTAAG TGATGCTTCA GTGACAGCTG 780
GAGGTTTCCA AATCAAATAT GTTGCAATGG ATCCTGTATC CAAATCCAGT CAAGGAAAAA 840
ATACAAGTAC TACTTCTACT GGAATAAAAA ACTTTTATAG TGGAAAGATT AGCCACTTAT 900
AAAAAAGGAA AAGGATGATC AAAACACACA GTGTTTATGT TGGAACTCTT TGGAACTCCT 960
TTGATCTCAC TGTATTATAT AACATTATAT TATTATTTT CTAAATGTGA AAGAAATACA 1020
TAATTTAGGG AAAATGTGAA AATATAGGAA ACTTTAAACG AGAAAAATGA ACCTCTCATA 1080
ATCCCACTGC ATAGAATAA CAAGCGTTAA CATTTTCATA TTTTTCCTT TCAGTCATTT 1140
TTGTATTTGT GGTATATGTA TATATGTACC TATATGTATT TGCATTGGA ATTTTGGAA 1200
CTGCTCTAT GTACAGTTT GTATTATCT TTTTAAATCT TGAACCTTAT GAACATTTTC 1260
TGAATCATT GATTATCTA CAAAACATG ATTTTAAACA GCTGTAAAT ATTCTATGAT 1320
ATGAATGTTT TATGATTTAT TTAAGCCTGT CTCTATTGTT GGAATTCAG GTCATTTCAT 1380
TAAATATTGT TGCAATAAAT ATCCTTCGGA ATTC
```

Seq ID NO: 507 Protein sequence
Protein Accession #: NP_009046.1

65
70

```
1 11 21 31 41 51
MIIILYLFLL LWEDTQGWGF KDGIFHNSIW LERAAGVYHR EARSQKYKLT YAEKAVCEP 60
EGGHLATYKQ LEAARKIGFH VCAAGWMAKG RVGYPIVKPG PNXFGKTGI IDYGI RLNRS 120
ERWDAYCYNP HAKECCGVPT DPKRIPKSPG PFNEYEDNQI CYWHIRLKYG QRIHLSFLDF 180
DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELFPDDIIS TGNVMTLKFL SDASVTAGGF 240
QIKYVAMPDV SKSSQGNKNTS TTSTGNKNFL AGRPSHL
```

75
80

Seq ID NO: 508 DNA sequence
Nucleic Acid Accession #: NM_001044.1
Coding sequence: 129..1991

85

```
1 11 21 31 41 51
ACCGCTCCGG AGCGGGAGGG GAGGCTTCGC GGAACGCTCT CGGCGCCAGG ACTCGGTCG 60
AAAGCCGAGG CCGGGGCGGC CAGACCAAGA GGAAGAAGAC ACAGAATTCC TCAACTCCCA 120
GTGTGCCCAT GAGTAAGAGC AATGCTCCG TGGGACTCAT GTCTTCGCTG GTGGCCCCGG 180
CTAAGGAGCC CAATGCCGTG GGCCCGAAGG AGGTGGAGCT CATCTTGTG AAGGAGCAGA 240
ACGGAGTGCA GCTCACAGC TCCACCTCA CCAACCCGCG GCAGAGCCCC GTGGAGGCC 300
AGGATCGGGA GACCTGGGGC AAGAAGATCG ACTTCTCCT GTCCGTCAAT GGCTTTGCTG 360
TGGACCTGGC CAACGCTGCG CGGTTCCCTT ACCTGTGCTA CAAAAATGGT GCGGTGCTCT 420
```

	TCCCTGGTCCC	CTACCTGCTC	TTTCATGGTCA	TTGCTGGGAT	GCCACTTTTC	TACATGGAGC	480
	TGGCCCTCGG	CCAGTTCAAC	AGGGAAGGGG	COGCTGGTGT	CTGGAAGATC	TGCCCCATAC	540
	TGAAAGGTGT	GGGCTTCAAG	GTCACTCTCA	TCTCACTGTA	TGTGGGCTTC	TTCTACAACG	600
	TCATCATCGC	CTGGGCGCTG	CACATCTCT	TCTCCTCCTT	CACCAAGGAG	CTCCCTCGGA	660
5	TCCACTGCAA	CAACTCTCTG	AACAGCCCCA	ACTGCTCGGA	TGCCATCTCT	GGTACTCCA	720
	GTGGAGACAG	CTCGGGCTCT	AACGACACTT	TTGGGACCAC	AOCTGCTGCC	GAGTACTTTG	780
	AACGTGGCGT	GCTGCACCTC	CACCAGAGCC	ATGGCATCGA	CGACCTGGGG	CCTCGCGGTG	840
	GGCAGCTCAC	AGCCTGCTGT	GTGCTGGTCA	TGCTGCTGCT	CTACTTCAGC	CTCTGGGAAG	900
	GGGTGAAGAC	CTCAGGGAAG	GTGGTATGGA	TCACAGCCAC	CATGCCATAC	GTGGTCTCTA	960
10	CTGCCCTGCT	CCTGCGTGGG	GTCAACCTCC	CTGGAGCCAT	AGACGGCATC	AGAGCATACC	1020
	TGAGCGTGA	CTTCTACCGG	CTCTGCGAGG	CGTCTGTTTG	GATTGACGGG	GCCACCCAGG	1080
	TGTGCTTCTC	CCTGGGCGTG	GGGTTCGGGG	TGCTGATCGC	CTTCTCCAGC	TACAACAAGT	1140
	TCACCAACAA	CTGCTACAGG	GACGCGATTG	TCACCACTCT	CATCAACTCC	CTGACGAGCT	1200
	TCTCCTCCGG	CTTCTGCTCT	TTCTCCTTCC	TGGGTACAT	GGCACAAG	CACAGTGTGC	1260
15	CCATCGGGGA	CGTGGCCCA	GACGGGCCAG	GGCTGATCTT	CATCATCTAC	CCGGAAGCCA	1320
	TGCGCAAGCT	CCCTCTGTCC	TCAGCCTGGG	COGTGGTCTT	CTTCATCATG	CTGCTCACC	1380
	TGGGTATCGA	CAGCGCCATG	GGTGGTATGG	AGTCAGTGAT	CACCGGGCTC	ATCGATGAGT	1440
	TCCAGCTGCT	GCACAGACAC	CGTGAGCTCT	TCACTGCTCT	CATCGTCTCT	GCGACCTTCC	1500
	TCCTGCTCCCT	CTTCTGCTCT	ACCAACGGTG	GCATCTAAGT	CTTCACTCTC	CTGGACCAAT	1560
20	TTGCAGCCGG	CACGTCCATC	CTCTTTGGAG	TGCTCATCGA	AGCCATCGGA	GTGGCCTGGT	1620
	TCTATGTTGT	TGGGCACTTG	AGCGACGACA	TCCAGCAGAT	GACCGGGCAG	CGGCCACGCC	1680
	TGTACTGGCG	GCTGTGCTGG	AAGCTGGTCA	GCCCTGCTCT	TCTCTGTTTC	GTGGTGGTGG	1740
	TCAGCATTTG	GACCTTCAGA	CCCCCCCACT	ACGGAGCCTA	CATCTTCCCC	GACTGGGCCA	1800
	ACGCGCTGGG	CTGGGCTCAT	GCCACATCCT	CCATGGCCAT	GGTGGCCATC	TATGCGGCTC	1860
25	ACAAGTTCTG	CAGCTGCTCT	GGGTCTTCTT	GAGAGAAACT	GGCCTAGGCC	ATTGCAACCG	1920
	AGAAGGACCG	TGAGCTGGTG	GACAGAGGGG	AGGTGCGCCA	GTTCACGCTC	CGCCACTGGC	1980
	TCAAGGTGTA	GAGGAGACAC	AGACGAAGAC	CCACGGAAGT	CATCTGCAAA	TGGGAGAGAC	2040
	ACGAACAAC	CAAGAAATC	TAAGTTTCTA	GAGAAAGGAG	GGCACTTCT	ACTCTTCAAC	2100
	CTCTACTGAA	AACACAACCA	ACAAAGCAGA	AGACTCTCT	CTTCTGACTG	TTTACACCTT	2160
30	TCCGTGCGGG	GAGCGCACCT	CGCGTGTCT	TGTGTGCTG	TAATAACGAC	GTAGATCTGT	2220
	GCAGCGAGGT	CCACCCCGTT	GTGTCTCTCT	CAGGGCAGAA	AAACGCTTAA	CTTCTGCTG	2280
	TCTGTGTGAG	GCTCTCTCC	TCCCTGCTCT	CTGCTCCCGG	CTCTGAGGCT	GCCCGAGGGG	2340
	CACGTGTGTT	TCAGCGGGG	ATCAGATCC	TTGTAGAGCG	AOCTGCTGAG	AATCCCGGTG	2400
	CTCACAGTAG	CTTCTAGAC	CATTACTTTT	GCCCATATTA	AAAAGCCAA	TGCTCTGCTT	2460
35	GGTTAGCTG	TGCAGAAAGT	GAAATGGAGG	AAACCACAAA	TTATGCAAAA	GTCCCTTCCC	2520
	GATGCGTGGC	TCCAGCAGA	GGCCGTAAAT	TGAGCGTTCA	GTGACACAT	TGCACACACA	2580
	GTCTGTTTCA	AGGCATTGGA	GGATGGGGT	CCTGGTATGT	CTCACCAGGA	AATTCTGTTT	2640
	ATGTTCTTGC	AGCAGAGAGA	AATAAACTC	CTTGAACCA	GCTCAGGCTA	CTGCCACTCA	2700
	GGCAGCCTGT	GGGTCTTGT	GGTGTAGGGA	ACGGCCTGAG	AGGAGCGTGT	CCTATCCCCG	2760
40	GACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	CCTATCCCCG	GACGCATGCA	GGGCCCCCAC	2820
	AGGAGCATGT	CCTATCCCCG	GACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	ACTATCCCCG	2880
	AACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	ACTATCCCCG	GACGCATGCA	GGGCCCCCAC	2940
	TGGAGCGTGT	ACTATCCCCG	GACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	CCTATCCCCG	3000
	GACCGGACGC	ATGCAGGGCC	CCCACAGGAG	CGTGTACTAC	CCCAGGACGC	ATGCAGGGCC	3060
45	CCCACAGGAG	CGTGTACTAC	CCCACAGGAG	ATGCAGGGCC	CCCACAGGAG	CGTGTACTAC	3120
	CCCAGGACGC	ATGCAGGGCC	CCCACAGGAG	CAGCCTGCGA	ACCAACATCT	TGCTGGGCTC	3180
	TGAGCGTGA	CTCCACGAA	GGGACCCAC	TGGAATTTTA	TTTCTCTCAG	GTGCGTGCCA	3240
	CATCAATAAC	AACAGTTTTT	ATGTTTGCAG	ATGGCTTTTT	AAAATCATAT	TTACCTGTGA	3300
	ATCAAAACAA	ATTCAAGAAAT	GCAGTATCCG	CGAGCCTGCT	TGCTGATATT	GCAGTTTTTG	3360
50	TTTACAGAA	TAATTAGCAA	TACTAGTGA	AGGATGTTGG	CCAAAGCTGT	CTTCTCATGG	3420
	CACACTGCCC	TCTGCCACTG	ACAGGAAAGT	GGATGCCATA	GTTTGAATTC	ATGCCCTCAAG	3480
	TGGGTGGGCC	TGCCTAGTGT	CTGCCGAGG	GCAGGGGCGG	TGCAGGGCCA	GTCTGGCTGT	3540
	TCCCTGCAA	GTGGACGTGG	GCTCCAGGGA	CTGGAGTGTG	ATGCTCGGTG	GGAGCGGTCA	3600
	GCCTGTGAAC	TGCCAGGACG	CTGCAGTTAG	CACAGAGGAT	GGCTTCCCCA	TTGCCCTCTG	3660
55	GGGAGGGACA	CAGAGGACGG	CTTCCCATC	GCCTTCTGCG	CGCTGCAGTC	AGCAGAGAGA	3720
	GGCGCTTCCC	CATTGCCTCT	TGGGGAGGGA	CACAGAGGAC	AGTTTCCCCA	TGCGCTTCTG	3780
	GTTGTTGAAG	ACAGCACAGA	GAGCGGCTTC	CCCATCGCCT	TCTGGGGAGG	GGCTCGGTGT	3840
	AGCAACCCAG	GTGTTGTCCG	TGCTGTGTGA	CCAATCTCTA	TTCAGCATCG	TGTGGGTCCC	3900
60	TAAGCAACAT	AAAAGACATC	CACAAATGAA	AAAAAAAAG	GAATTC		

Seq ID NO: 509 Protein sequence
Protein Accession #: NP_001035.1

	1	11	21	31	41	51	
65	MSKSKCSVGL	MSSVVAPEKE	PNAVGPKEVE	LILVKEQNGV	QLTSSTLTNP	RQSPVEAQDR	60
	ETWKKIDFL	LSVIGPAVDL	ANVWRFPYLC	YKNGGGAFLV	PYLLEFVIA	MPLFYMELAL	120
	GQFNREGAAG	VWKICPILKG	VGFTVILISL	YVGFPPNVII	AWALHYLFSS	FTTELFWIHC	180
	NNSWNFPNCS	DAHFGDSSGD	SSGLNDTFGT	TPAAEYPERG	VLHLHQSHGI	DDLGPFRWQL	240
70	TACLVLVIVL	LYFSLWKGVK	TSGKVVWITA	TMPIVVLTL	LLRGVTLPGA	IDGIRAYLSV	300
	DFYRLCEASV	WIDAAQTQCF	SLGVGFGVLI	AFSSYNKFTN	NCYRDAIVTT	SINSLTSFSS	360
	GFVVFSPLGY	MAQKHSVPIG	DVAKDGPGLI	PIIYPEAIAT	LPLSSANAVV	FFIMLLTLGI	420
	DSAMGGMESV	ITGLIDEFQL	LHRHRELFTL	FIVLATFLLS	LFCVTNGGIY	VFTLLDHFAA	480
	GTSILFGVLI	EAIGVAFVYG	WQFSDDIQQ	MTGQRPSLYW	RLCNKLVSPC	FLLFVVVVISI	540
75	VTFRPPHYGA	YIFPDWANAL	GWVIATSSMA	MVPIYAAKYF	CSLPGSFREK	LAYAIAPKED	600
	RELVDREVR	QTLRHLWLV					

Seq ID NO: 510 DNA sequence
Nucleic Acid Accession #: NM_001216.1
Coding sequence: 43..1422

	1	11	21	31	41	51	
85	GCCCGTACAC	ACCGTGTGCT	GGGACACCCC	ACAGTCAGCC	GCACTGGCTCC	CCTGTGCCCC	60
	AGCCCTGGC	TCCCTCTGTT	GATCCCGGCC	CCTGCTCCAG	GCCTCACTGT	GCAACTGCTG	120
	CTGTCACTGC	TGCTTCTGAT	GCCTGTCCAT	CCCCAGAGGT	TGCCCGGAT	GCAGAGGAT	180
	TCCCTCTTGG	GAGGAGGCTC	TTCTGGGGAA	GATGACCCAC	TGGGCGAGGA	GGATCTGCC	240

	AGTGAAGAGG	ATTACCCAG	AGAGGAGGAT	CCACCCGGAG	AGGAGGATCT	ACCTGGAGAG	300
	GAGGATCTAC	CTGGAGAGGA	GGATCTACCT	GAAGTTAAGC	CTAAATCAGA	AGAAGAGGGC	360
	TCCTGAACT	TAGAGGATCT	ACCTACTGTT	GAGGCTCCTG	GAGATCTCTCA	AGAACCCGAC	420
5	AATAATGCC	ACAGGGGACAA	AGAAGGGGAT	GACCAAGATC	ATTGGCGCTA	TGGAGGGGAC	480
	CCGCCCTGGC	CCCGGGTGTG	CCCAGCCTGC	GCGGGCCGCT	TCCAGTCCCC	GGTGGATATC	540
	CGCCCCCAGC	TCSCCGCCTT	CTGCCCGGCC	CTGCCCGCCC	TGGAATCCTT	GGCTTCCAG	600
	CTCCCCCGGC	TCCCAGAACT	GCGCCTGCGC	AACAATGGCC	ACAGTGTGCA	ACTGACCCCTG	660
	CCTCCTGGGC	TAGAGATGGC	TCTGGGTCCC	GCGCGGGAGT	ACCGGCTCTT	GCAGCTGCAT	720
10	CTGCATGGG	GGGCTGCAGG	TCGTCCGGGC	TCGGAGCACA	CTGTGGAAGG	CCACCGTTTC	780
	CCTGCCGAGA	TCCACGTGGT	TCACCTCAGC	ACCGCCTTTG	CCAGAGTTGA	CGAGGCCTTG	840
	GGGCGCCCCG	GAGGCTTGGC	CGTGTGGGCC	GCCTTTCTGG	AGGAGGGCCC	GGAGAAAAC	900
	AGTGCCTATG	AGCAGTTGCT	GTCTCGCTTG	GAAGAAATCG	CTGAGGAAGG	CTCAGAGACT	960
	CAGGTCCAGT	GCTGACAT	ATCTGCACTC	CTGCCCTCTG	ACTTCAGCCG	CTACTTCCAA	1020
15	TATGAGGGT	CTCTGACTAC	ACCGCCCTGT	GCCAGGGTG	TCATCTGGAC	TGTGTTTAA	1080
	CAGACAGTGA	TGCTGAGTGC	TAAGCAGCTC	CACACCCCTT	CTGACACCCT	GTGGGGACCT	1140
	GGTGACTCTC	GGCTACAGCT	GAACCTTCGA	GCGAGCAGC	CTTTGAATGG	GCGAGTGATT	1200
	GAGGCTCTCT	TCCCTGCTGG	AGTGGACAGC	AGTCTCTGGG	CTGCTGAGCC	AGTCCAGCTG	1260
	AAATCTCTGC	TGGCTGCTGG	TCACATCTTA	GCCCTGGTTT	TTGGCCTCCT	TTTGTCTGTC	1320
20	ACCAGCTGG	CGTCTCTTGT	GCAGATGAGA	AGGCAGCACA	GAAGGGGAAC	CAAAGGGGGT	1380
	GTGAGCTACC	GGCCAGCAGA	GGTAGCCGAG	ACTGGAGCCT	AGAGGCTGGA	TCTTGGAGAA	1440
	TGTGAGAAGC	CAGCCAGAGG	CATCTGAGGG	GGAGCCGGTA	ACTGTCTCTG	CCTGCTCATT	1500
	ATGCCACTTC	CTTTTAACTG	CCAAGAAATT	TTTTAAATA	AATATTATA	AT	

Seq ID NO: 511 Protein sequence
Protein Accession #: NP_001207.1

	1	11	21	31	41	51	
	MAPLCPSFWL	PLLIPAPAPG	LTVQLLSLL	LLMPVHPQRL	PRMQEDSPLG	GGSSGEDDPL	60
30	GEEDLPSEED	SPREEDPPGE	EDLPGEEDLP	GEEDLPVVKP	KSEEEGSLKL	EDLPTEAPG	120
	DPQSPQNNAH	RDKEGDDQSH	WRYGGDPPWP	RVSPACAGRF	QSPVDIRFQL	AAFCPALRPL	180
	ELLGFQLPPL	PELRLRNHNG	SVQLTLPPGL	EMALPGREY	RALQLHLHWG	AAGRPGSEHT	240
	VEGHRPABEI	HVVHLSTAF	RVDEALGRPG	GLAVLAFL	EGPEENSAYE	QLLSRLLEIA	300
	EEGSETQVPG	LDISALLPSD	FSRYPYEBS	LITPPCAGGV	IWTFNQITVM	LSAKQLHTLS	360
35	DTLMGPGDSR	LQLNFRATQP	LANGRVIEASF	PAGVDSPPRA	AEPVQLNSCL	AAGDILALVF	420
	GLLPATSV	FLVQMRQRH	RGTKGGVSYR	PAEVAETGA			

Seq ID NO: 512 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..3978

	1	11	21	31	41	51	
	ATGGTGGGTG	AAGGACCCTA	CCTTATCTCA	GATCTGGACC	AGCGAGGCCG	GCGGAGATCC	60
45	TTTGAGAAA	GATATGACCC	CAGCCTGAAG	ACCATGATCC	CAGTGCGACC	CTGTGCAAGG	120
	TTAGCACCCA	ACCCGGTGA	TGATGCCGGG	CTACTCTCCT	TGCCACACAT	TTCTGGCTC	180
	ACGCGGTGA	TGGTGAAAGG	CTACCGGCAA	AGGCTGACCG	TAGACACCCCT	GCCCCATTG	240
	TCGACATATG	ACTCATCTGA	CACCAATGCC	AAAAGATTTC	GAGTCTCTTG	GGATGAAGAG	300
	GTAGCAGGGT	TGGTCTCTGA	GAAGGCCCTCT	CTGAGCCACG	TGGTGTGGAA	ATTCCAGAGG	360
50	ACACGCGTGT	TGATGGACAT	CGTGGCCAAC	ATCCTGTGCA	TCATCATGCG	AGCCATAGGG	420
	CCGACAGTTC	TCATTACCA	AATCCTCCAG	CAGACTGAGA	GGACCTCTGG	GAAAGTCTGG	480
	GTTGGCATTG	GACTGTGCAT	AGCCCTTTTT	GCCACCGAGT	TTACCAAAAT	CTTCTTTTGG	540
	GCCCTTGCCCT	GGGCCATCAA	CTACCGCAGG	GCCATCCGGT	TGAAGGTGGC	GCTCTCCACC	600
	TTGGTTTTTG	AAAACCTAGT	GTCCCTCAAG	ACATTGACCC	ACATCTCTGT	TGGCGAGGTG	660
55	CTCAATATAC	TGTCAGATGA	TAGCTATTCT	TTGTTTGAAG	CTGCCCTTGT	TTGTCTTTTG	720
	CCAGCCACCA	TCCGATCTCT	AATGGTCTTT	TGTGCGGGGT	ACGCCCTTTT	CATTCTGGGG	780
	CCCAAGCTC	TCATCGGGAT	ATCAGTGTAT	GTCAATTCA	TACCCGTCCA	GATGTTTATG	840
	GCCAAGCTCA	ATTGAGCTTT	CCGAAGGTCA	GCAATTTTGG	TGACAGACAA	GCGAGTTCAG	900
60	ACAATGAATG	AGTTTCTGAC	CTGCATCAGG	CTGATCAAAA	TGTATGCCCTG	GGAGAAATCT	960
	TTTACCAACA	CTATCCAGAA	TATAAGAAAG	AGGGAAGAA	AATTACTGGA	AAAAGCTGGA	1020
	TTTGTCCAAA	GTGGAATCT	TGCCCTGGCC	CCCATCTGTT	CCACCATAGC	CATCGTGTCTG	1080
	ACATTATCTT	GCCATCTCT	CCTGAGACGC	AAACTCACCG	CACCGGTGGC	ATTAGTGTG	1140
	ATTGCCATGT	TTAATGTAAT	GAAATTTTCC	ATTGCAATCT	TGCCCTTCTC	CATCAAAGCA	1200
	ATGGCTGAAG	CGAATGTCTC	TCTAAGGAGA	ATGAAGAAAA	TTCTCATAGA	TAAAGCCCCC	1260
65	CCATCTTACA	TCACCCAAAC	AGAAGACCCA	GATACTGTCT	TGCTTTTAGC	AAATGCCACC	1320
	TTGACATGGG	AGCATGAAGC	CAGCAGGAAA	AGTACCCCAA	AGAAATTGCA	GAACAGAAA	1380
	AGGCATTAT	GCAAGAAACA	GAGGTCAGAG	GCATACAGTG	AGAGGAGTCC	ACCAGCCAG	1440
	GGAGCCACTG	GCCAGAGGA	GCAAAGTGAC	AGCCTCAAAT	CGTTCTGCA	CAGCATAAGC	1500
	TTGTGTGTGA	GAAAGTTATG	TCGTTATCCC	GAAGCCCAGC	TCCTGGCTTG	GAGGTGGCCA	1560
70	GCAGTGTTTG	TGGGAGAAT	CATCAGAGGA	TACAGGCCCTC	ATGGATTTTC	TGCTAAAGAC	1620
	AAGGATGAAT	CTAGAAGGCT	TCTTACTTGG	CCCCAAGAG	TGGATAGGAC	TCAAAGGGCA	1680
	GCCAAATACC	TGGGGAAGAT	CTTGGGAATA	TGTGGGAATG	TGGGAAGTGG	AAAGAGCTCC	1740
	CTCCTTGCG	CTCTCTAGG	ACAGATGCAG	CTGCAGAAAG	GGGTGGTGGC	AGTCAATGGA	1800
	ACTTTGGCCT	ACGTTTCACA	GCAGGCATGG	ATCTTTTCATG	GAAATGTGAG	AGAAAACATA	1860
75	CTCTTTGGAG	AAAAGTATGA	TCACCAAGG	TATCAGCACA	CAGTCCGCGT	CTGTGGCCTC	1920
	CAGAAGGACC	TGAGCAACCT	CCCCTATGGA	GACCTGACTG	AGATTGGGGA	GCGGGGCCCTC	1980
	AACCTCTCTG	GGGGGCAGAG	GCAGAGGATT	AGCCTGGCCC	GCGCTGTCTA	CTCGACCGT	2040
	CAGCTCTACC	TGCTGGACGA	CCCCCTGTGG	GCGGTGGACG	CCCACGTGGG	GAAGCACGTC	2100
	TTTGAGGAGT	GCATTAAGAA	GACGCTCAGG	GGAAAGACAG	TGCTCTCTGT	GACCCACGAC	2160
80	CTACAGTTCT	TAGGTCTCTG	TGATGAAGTT	ATTTTATTAG	AAGATGGAGA	GATTTGTGAA	2220
	AAGGGAACCC	ACAAGGAGTT	AATGGAGGAG	AGAGGGCGCT	ATGCAAAACT	GATTCAACAC	2280
	CTGCGAGGAT	TGCAGTTCAA	GGATCCTGAA	CACCTTTACA	ATGCAGCAAT	GTTGGAAGCC	2340
	TTCAAGGAGA	GCCCTGCTGA	GAGAGAGGAA	GATGCTGGTA	TAATCGGGTA	CCTCTTTTCT	2400
	CTCTTCACTG	TGTTCTCTTT	CCTCTGATG	ATTGGCAGCG	CTGCCTTCAG	CAACTGGTGG	2460
85	CTGGGTCTCT	GGTTGGACAA	GGGCTCACGG	ATGACCTGTG	GGCCCCAGGG	CAACAGGACC	2520
	ATGTGTGAGG	TGGCGCGGT	GCTGGCAGAC	ATCGTTCAGC	ATGTGTACCA	GTGGGTGTAC	2580
	ACTGCAAGCA	TGGTGTTCAT	GCTGGTGTTC	GGGCTCACCA	AAGGCTTCGT	CTTCACCAAG	2640

5
10
15
20
25

```

ACCACTGTA TGGCATCTC CTCTCTGCAT GACACGCTGT TTGATAAGAT CTTAAGAGC 2700
CCRAATGAGT TCTTTGACAC GACTCCCACT GGCAGGCTAA TGAACCGTTT TTCCAAGGAT 2760
ATGGACGAGC TGGAGTGTAG GCTGCCGTTT CACGACAGAGA ACTTTCTGCA GCAGTTTTTT 2820
ATGGTGGTGT TTAATCTCGT GATCTTGGCT GCTGTGTTTC CTGCTGTCCCT TTAGTCTGTG 2880
GCCAGCCTTG CTGTAGGCTT CTTCATTCTG TTACGCATTT TCCACAGAGG AGTCCAGGAG 2940
CTCAAGAGAG TGGAGAAATG CAGCCGGTCA CCCTGGTTCA CCCACATCAC CTCTCCCATG 3000
CAGGGCCTGG GCATCATTTA CGCCTATGGC AAGAAGGAGA GCTGCATCAC CTATACITCA 3060
TCCAAAGGCC TGTTCATTGT ATACATCATC CAGCTGAGCG GACTGCTCCA AGTGTGTGTG 3120
CGAAACGGGA CAGAGACGCA AGCCAAATTC ACCTCCGTGG AGCTGCTCAG GGAATACATT 3180
TCGACCTGTG TTCTGAATG CACTCATCCC CTCAAAGTGG GGACCTGTCC CAAGGACTGG 3240
CCCAGCTGTG GGGAGATCAC CTTTCAAGAC TATCAGATGA GATACAGAGA CAACACCCCC 3300
CTTGTCTCG ACAGCCTGAA CTTGAACATA CAAAGTGGGC AGACAGTCGG GATTGTGGA 3360
AGAACAGGTT CCGGAAAGT ATCGTTAGGA ATGGCTTTGT TTGCTGTGGT GGAGCCAGCC 3420
AGTGGACCAA TCTTTATTGA TGAGGTGGAT ATCTGCATTC TCAGCTTGGA AGACCTCAGA 3480
ACCAAGCTGA CTGTGATCCC ACAGGATCCT GTCTGTGTTG TAGGTACAGT AAGGTACAA 3540
TTGGATCCCT TTGAGAGTCA CACCGATGAG ATGCTCTGGC AGGTTCTGGA GAGAACATTC 3600
ATGAGAGACA CAATATGAA ACTCCCAGAA AAATTACAG CAGAAGTCAC AGAAATGGA 3660
GAAACTTCT CAGTAGGGGA ACCTCAGCTG CTTTGTGTGG CCGAGCTCT TCTCCGTAAT 3720
TCAAAGATCA TTCTCCTTGA TGAAGCCACC GCCTCTATGG ACTCCAAGAC TGACACCCCTG 3780
GTTCAAGACA CCAATCAAGA TGCCTTCAAG GGCTGCACCT TGCTGACCAT CGCCACCCGC 3840
CTCAACACAG TTCTCAACTG CGATCAAGTC CTGTTTATGG AAAATGGGAA GGTGATTGAG 3900
TTTGACAAGC CTGAAGTCTG TGCAGAGAAG CCAGATTCTG CATTTCGCGT GTTACTAGCA 3960
GCAGAAGTCA GATTGTAG

```

Seq ID NO: 513 Protein sequence
Protein Accession #: Eos sequence

30
35
40
45
50

```

1 11 21 31 41 51
MVGEOPYLIS DLDQRRRRS FAERYDPSLK TMIPVRPCAR LAPNFVDDAG LLSFATPSWL 60
TPVMVKYRQ RLTVDTLPL STYDSSDTNA KRFRVLWDES VARVGPEKAS LSHVVVKFQR 120
TRVLMDIVAN ILCIIMAAIG PTVLIHQILQ QTERTSGKVV VGIGLCIALF ATEPTKVFPW 180
ALAWAINYRT AIRKLVALST LVFENLVSEK TLTHISVGEV LNILSSDSYS LFEALFCPL 240
PATIPILMFV CAAYAFFILG PTALIGISVY VIPIPVQMFM AKLNSAFRRS AILVTDKRVQ 300
TMEFLTICIR LIKMYAWEKS PTNTIQDIRR RERKLLEKAG FVQSGNSALA PIVSTIAIVL 360
TLSCHILLRR KLTAIVAFSV IAFNVVMKFS IAILPFSSKA MAEAMVSLRR MKKILIDKSP 420
PSYITQPEDE DTVLLANAT LTWEHEASRK STPKKLQKQK RHLCKKQRE AYSESPPAK 480
GATQPEBQSD SLKSVLHSIS PVVRKLCRYP EAQLLAWRWP AVFVGRIIRG YRPHGFSKAD 540
KDESRLLTW PQEVDRTORA AKYLGKILGI CGNVGSGKSS LLAALLGQM QLVGVVAVNG 600
TLAYVQQAW LFHGNVRENI LPGEKYDHQR YQHTVRVCG LKDLNLPYG DLTEIGERGL 660
NLSGGQRRI SLARAVYSR QLYLLDDPLS AVDAHVGKHV FEECIKKTLR GKTVVLVTHQ 720
LQFLESCEDE ILLDEGEICE KGHKELMEE RGRYAKLIHN LRGLQFKDPE HLYNAAMVEA 780
FKESPAEREE DAGIIGYLLS LFTVFLFLM IGSAAPSNWW LGLWLDKGRS MTCGPQGNRT 840
MCEVGAVLAD IGQHVYQWVY TASMVFMVLF GVTKGFTVTK TLMASSSLH DTVFDKILKS 900
PMSFFDTTPT GRIMNRFSD MDELDRLPF HAENFLQQFF MVVFILVILA AVPPAVLLVV 960
ASLAVGFPLI LRIFHRGVQE LKKVENVSRS PWFTHITSSM QGLGIIHAYG KKECITYTS 1020
SKGLSLSYII QLSGLQVCV RTGTETQAKF TSVELLREYI STCVPECTHP LKVGTCPKDW 1080
PSCGEITFRD YQMYRDRNTF LVLDLNLINI QSGQTVGIVG RTGSGKSSLG MALFRLVEPA 1140
SGTIFIDEVD ICILSLLEDR TKLTVPQDP VLFVGTVRYN LDPFESHTDE MLWQVLERTF 1200
MRDTIMKLEP KLQAEVTENG ENFSVGERQL LCVARALLRN SKIILLDEAT ASMDSKTDTL 1260
VQNTIKDAFK GCTVLTIAHR LNTVLNCDHV LVMENKRVIE FDKPEVLARK PDSAFAMILLA 1320
AEVRL

```

Seq ID NO: 514 DNA sequence
Nucleic Acid Accession #: Z31560
Coding sequence: 1-966

60
65
70
75
80

```

1 11 21 31 41 51
CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCGCGG CCCGACGAA 60
ACTTCGGGGG GCGGCGCGGG CAACTCCACC GCGGCGCGGG CCGGCGGCAA CCAGAAAAAC 120
AGCCCGGACC GCGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGTTCCCG CCGGCGCGGG 180
CGCAAGATGG CCCAAGAGAA CCCCAAGATG CACAACTCGG AGATCAGCAA GCGCCTGGGC 240
GCGGAGTGA AACTTTTGTG GGAGACGGAG AAGCGGCGGT TCATCGACGA GGCTAAGCGG 300
CTGCGAGCGC TGCACATGAA GGAGCACCGG GATTATAAAT ACCGGCCCCG GCGGAAAAAC 360
AAGACGCTCA TGAAGAAGGA TAAGTACACG CTGCCCGGCG GGCTGCTGGC CCGGCGCGGC 420
AATAGCATGG CGAGCGGGGT CCGGGTGGGC GCGGCGCTGG GCGCGGCGGT GAACACGCGC 480
ATGGACAGTT ACGGCGACAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540
CAGCTGGGCT ACCCGCAGCA CCGCGGCTCT AATGCGCACG GCGCAGCGCA GATGCAGCCC 600
ATGCACGCT ACGACGTGAG CGCCCTGCAG TACAACTCCA TGACCACTCT GCAGACCTAC 660
ATGAACGGCT CGCCCACTTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT 720
CTTGGCTCCA TGGTTCGGT GGTCAAGTCC GAGGCGAGCT CCAGCCCCCC TGTGTTTACC 780
TCTTCTCTCC ACTCCAGGGC GCGCTGCCAG GCGGCGGACC TCGGGACAT GATCAGCATG 840
TATCTCCCG GCGCGAGGT GCGGGAACCC GCGGCCCCCA GCAGACTTCA CATGTCCAG 900
CACTACACGA GCGGCGCGGT GCGGCGCACG GCCATTAAAG GCACACTGCC CCTCTCACAC 960
ATGTGAGGGC CGGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAAGAAAA ACGAGGGAAA 1020
TGGGAGGGGT GCAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCCG TACGCTCAA 1080
AAAAA

```

Seq ID NO: 515 Protein sequence
Protein Accession #: CAA83435

85

```

1 11 21 31 41 51
HSARMYMMME TELKPPGPQQ TSGGGCGNST AAAAGGNQKN SPDRVKRFMN AFMVNSRGQR 60
RMAQENPKM HNSEISKRLG AEWKLLSETE KRPFIDEAKR LRALHMKHP DYKYRPRKT 120
KTLMKDKKIT LPGGLLAPGG NSMASGVGVG AGLGAGVNQR MDSYAHMNGV SNGSYMMQD 180

```

QLGYPHPLGL NAHGAAQMQP MHRVDVSALQ YNSMTSSQTY MNGSPYSMS YSQQGTGMA 240
LGSMSGVVKS EASSSPVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPVEP AAPSLHMSQ 300
HYQSGVPVGT AINGTLELSH M

5 Seq ID NO: 516 DNA sequence
Nucleic Acid Accession #: U91618
Coding sequence: 29..541

10 1 11 21 31 41 51
CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60
CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAAATGAA 120
AGCATTAGAA GCAGATTTC TGAACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180
15 TCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC 240
AGCTGAGGAA ACAGGAGAAG TTCATGAAGA GGAGCTTGTG GCAAGAAGGA AACTTCCTAC 300
TGCTTAGATG GGCTTTAGCT TGAAGCAAT GTTGACAATA TACCAGCTCC ACAAAATCTG 360
TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA 420
TGACAAAAT GGAAGGAAG AAGTCATAAA GAGAAAAATT CCTTATATTC TGAACCGCA 480
20 GCTGTATGAG AATAAACCCA GAAGACCCCTA CATACTCAA AGAGATTCTT ACTATTACTG 540
AGAGATAA TCAATTATTT ACATGTGATT GTGATTCATC ATCCCTTAAT TAAATATCAA 600
ATTATATTTG TGTGAAATG TGACAAACAC ACTTATCTGT CTCTTCTACA ATTGTGGTTT 660
ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAA TAAATCTAAA 720
TCTTCAAAA AAAAAAAA AAATGGGGCC GCAATT

25 Seq ID NO: 517 Protein sequence
Protein Accession #: AAB50564

30 1 11 21 31 41 51
MMAGMKIQLV CMLLLAPSSW SLCSDEEEM KALRADPLTN MHTSKISKAH VPSWKMTLLN 60
VCSLVNINLS PAETGTGVHE EELVARRKLP TALDGFSLA MLTIYQLHKI CHSRAFQHWE 120
LIQEDILDG NDKNGKEVI KRKIPYILKR QLYENKPRP YILKRDSSYY

35 Seq ID NO: 518 DNA sequence
Nucleic Acid Accession #: NM_006536.2
Coding sequence: 109..2940

40 1 11 21 31 41 51
ACCTAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60
ATGTATGACG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120
AGCATTGCGAG TCTCTATTGG CAACCTGAAG TTGTGACTC TCCTGGTTGC CTTAAGTTCA 180
GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC 240
ATTGCAATTA ATCTCAGAGT ACCTGAGAAT CAGAACCTCA TCTCAACAT TAAGGAAATG 300
45 ATAACTGAAG CTTCAATTTA CCTATTAAAT GCTACCAAGA GAAGAGTATT TTTGAGAAAT 360
ATAAAGATT TAAATACCTGC CACATGGAAG GCTAATAATA ACAGCAAAAT AAAACAAGAA 420
TCATATGAAA AGGCAAAATGT CATAGTGACT GACTGGTATG GGGCAGATGG AGATGATCCA 480
TACACCTTAC AATACAGAGG GTGTGAAAA GAGGAAAAAT ACATTCAATT CACACCTAAT 540
TTCTACTGCA ATGATAACT AACAGCTGGC TACGGATCAC GAGGCGGAGT GTTTGTCCAT 600
50 GAATGGGCC ACCCTCGTGG GGGTGTGTTG GATGAGTATA ACAATGACAA ACCTTTCTAC 660
ATAAATGGGC AAAATCAAT TAAAGTGACA AGGTGTTTAT CTGACATCAC AGGCATTTTT 720
GTGTGTGAAA AAGGTCTTGG CCCCAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780
GGATGCACTT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840
AGTTTATCTT CTGTGGTTGA ATTTTGTAAAT GCAAGTACCC ACAACCAAGA AGCACCACAA 900
55 CTACAGAAAC AGATGTGCGAG CCTCAGAACT GCATGGGATG TAATCACAGA CTCCTGCTGAC 960
TTTCACACA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCTCCACAT ATTCTGCTT 1020
GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTCTGGATG TGTCCAGCAA GATGGCAGAG 1080
GCTGACAGAC TGTCTCAACT ACAACAAGCC GCAGAAATTT ATTTGATGCA GATTGTTGAA 1140
ATTCTACCT TGTGGGCTG TGCCAGTTTC GACAGCAAG GAGAGATCAG AGCCAGCCTA 1200
60 CACCAATTA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260
TCAGCTAAAA CAGACATCAG CATTTGTTCA GGGCTTAAGA AAGGATTGGA GGTGGTTGAA 1320
AACTGAATG GAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCG AGATGATAAG 1380
CTTCTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCATTGCC 1440
65 CTGGGTTTAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500
TTCTTTGTT CAGATATATC AAATCTCAAT AGCATGATTG ATGCTTTCAG TAGAATTTC 1560
TCTGGAACCT GAGACATTTT CCAGCAACAT ATTCAGCTTG AAGTACAGG TGAATAATGTC 1620
AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680
ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCTGAGA TTATATTATT TGATCCTGAT 1740
70 GGACGAAAT ACTACACAA TAAATTTATC ACCAATCTAA CTTTTGGGAC AGCTAGTCTT 1800
TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCTGAAACAA TACCATCAT 1860
TCTCTGCAAG CCCTGAAAGT GACAGTGACC TCTCGCGCT CCACTCAGC TGTGCCCCCA 1920
GCCACTGTGG AAGCCTTTGT GGAAGAGAC AGCCTCCATT TTCTCATCC TGTGATGATT 1980
TATGCCAATG TGAACAGGG ATTTTATCCC ATTCTTAATG CCAGTGTAC TGCCACAGTT 2040
75 GAGCCAGAGA CTGGAGATCC TGTACGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100
GTTATAAAAA ATGATGGAAT TTACTCGAGG TATTTTCTCT CCTTGTCTG AAATGGTAGA 2160
TATAGCTTGA AAGTGATGT CAATCACTCT CCCAGCATAA GCACCCAGC CCACCTCTATT 2220
CCAGGAGTCT ATGCTATGTA TGTACCAAGT TACACAGCAA AOGGTAATAT TCAGATGAAT 2280
GCTCCAGAGA AATCAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGCTT TAGCCGAGTC 2340
80 AGCTCAGGAG GCTCCTTTTC AGTGCTGGGA GTTCCAGCTG GCCCCCACC TGTGTGTTT 2400
CCACCATGCA AATTTATTGA CCTGGAAGCT GTAAAGTAG AAGAGGAATT GACCCTATCT 2460
TGGACAGCAC CTGGAAGA CTTTGATCAG GGCAGGCTA CAGCTATGA AATAAGAAATG 2520
AGTAAGATG TACAGAAAT CCAAGATGAC TTTAACCAATG CTATTTTAGT AATATACATCA 2580
AAGCGAAATC CTGCAAGC TGGCATCAGG GAGATATTTA CGTTCTCACC CCAGATTTC 2640
85 ACGAATGGAC CTGAACATCA GCCAAATGGA GAAACACATG AAAGCCACAG AATTTATGTT 2700
GCAATACGAG CAATGGAATG GAATCTCTTA CAGTCTGCTG TATCTAACAT TGCCAGGCG 2760
CCTCTGTTTA TTCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820
GGAGTTTATA CAGCAATGGG TTTGATAGGA ATCAATTGCC TTATTATAGT TGTGACACAT 2880

5 CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATTA 2940
 ATAAATATCC AAAGTGTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAAA 3000
 CATACTAACA AAGTCAAAAT AACATCAAAA CTGTATTAAA ATGCAITGAG TTTTGTGACA 3060
 ATACAGATAA GATTTTACCA TGGTAGATCA ACAATCTCTT TTGGGGGTAG ATTAGAAAAC 3120
 CCTTACACTT TGCCTATGAA CAAATAATAA AAATTATTCT TTAAGTAAT GTCTTTAAAG 3180
 GCAAAGGGAA GGTAAAGTC GGACCAAGT CAAGGAAAGT TTGTTTTATT GAGGTGGAAA 3240
 AATAGCCCAA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA 3300
 TCATTAGTAT ACTTGTATTA ATTTTCTTT TCTCCTTATC TGTGCAGTAC AGGTTGCTTG 3360
 TTTACATGAA GATCATGCTA TATTTATAT ATGTAGCCCC TAATGCAAG CTCTTTACCT 3420
 10 CTGTCTATT TGTATATAT ATTTACAGAT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480
 TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540
 TTTATGACAA AGGTCTATTG AATTATTGTT TGTGTAAGTT TCTACTCCCA TCAAAGCAGC 3600
 TTTCTAAGTT TATTGCTCTG GGTATTATTG GAATGATAGT TATAGCCCCN TATAATGCCT 3660
 15 TACCTAGGAA A

Seq ID NO: 519 Protein sequence
 Protein Accession #: NP_066527.1

20 1 11 21 31 41 51
 | | | | |
 MTQRSIAGPI CNLKFTVLIV ALSSELPFLG AGVQLQDNGY NGLLIINPQ VPENQNLIEN 60
 IKEMITEASF YLFNAIKRRV FFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAAH 120
 GDDPYTLQYR GCKEKGKYYH PTPNPLNDN LTAGYGSRRG VEVHEWAHLR WGVFDEYNND 180
 25 KPFYINGNQ IKVTRCSSDI TGIFVCEKGP CPQENCIISK LFKEGCTFIY NSTQNTASI 240
 MFMQSLSSVV EFCNASTHNQ EAPNLQNMOC SLRSAMDVIT DSADPHHSFP MNGTELPPPP 300
 TFSLVQAGDK VVCLVLVDVSS KMAEADRLLO LQQAEEFYIM QIVEIHTFVG IASFDSKGEI 360
 RAQLHQINSN DDKRLKLVSYL PTTVSAKTDI SICSGLKKGK EVVEKLGKKA YGSMILVTS 420
 GDDKLLGNCL PTVLSSGSTI HSIALGSSAA PNLEELSRLT GGLKFPVPDI SNSNSMIDAF 480
 30 SRISSTGDI FQHQIQLST GENVKPHQL KNTVTVDNTV GNDTMFLVTV QASGPPPEIL 540
 FDPDGRKYYT NNPITNLTFR TASLWIPGTA KPGHWYITLN NTHSLQALK VVTSRASNS 600
 AVPPATVEAF VERDSLHFPF PVMYIANVKQ GFYPILNATV TATVEPETGD PVTLLRLDDG 660
 AGADVINKDG IYSRYFFSFA ANGRYSLKVH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720
 IQMNAPRKSV GRNEERKMG PSRVSSGGSP SVLGVPAGPH PDVFPCKII DLEAVKVEEE 780
 35 LTLSWTAPGE DFDQGGATSY BIRMSKSLQI IQDDFNNAIL VNTSKRNPQ AGIREIFTFS 840
 PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900
 LILKGVLTAM GLIGIICLLI VVTHHTLSRK KRADKENGTT KLL

Seq ID NO: 520 DNA sequence
 Nucleic Acid Accession #: NM_000228.1
 Coding sequence: 82..3600

40 1 11 21 31 41 51
 | | | | |
 GCITTCAGCG GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAAGG TCCTTTCTGG 60
 GGATCACCCC ATTGGCTGAA GATGAGACCA TTCTTCTCT TGTGTTTTGC CCTGCCTGGC 120
 CTCTGCTGAT CCCAACAAAG CTGCTCCCGT GGGGCGCTGT ATCCACCTGT TGGGCACTGT 180
 CTGTGTGGGA GAGCCCGGT TCTCCGAGCT TCATCTACCT GTGGACTGAC CAAGCCTGAG 240
 45 ACCTACTGCA CCCAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCCAGGCAG 300
 CCTCACACT ACTACAGTCA CCGAGTAGAG AATGTGGCTT CATCTCCGG CCCCATGCGC 360
 50 TGGTGGCAGT CCAGAAATGA TGTGAACCCCT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420
 TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGCCCGCCGG CATGCTGATT 480
 GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC 540
 ACCTCCACCT TCCTCGGGT CCGCCAGGT CGGCCCTCAGA GCTGGCAGGA TGTTCGGTGC 600
 CAGTCCCTGC CTGAGAGGCC TAATGCAAGC CTAATGGGG GGAAGGTCCA ACTTAACCTT 660
 55 ATGGAATTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAAA AAATTCAAGA GGTGGGGGAG 720
 ATCACAACAT TGAGAGTCAA TTTCAACAGG CTGCGCCCTG TGCCCAAGG GGGCTACCA 780
 CCTCCACGCG CTACTATATG TGTGTCCAG CTCTCTCTGC AGGGGAGCTG CTCTCTCTGC 840
 GGGCATGCTG ATGCTGCTGC ACCCAAGCCT GGGGCGCTGT CAGGCCCTC CACCGCTGTG 900
 60 CAGGTCCAGT ATGTCTGTGT CTGCCAGCAC AACACTCGCG GCCCAATTG TGAGCGCTGT 960
 GCACCTCTCT ACACCAACCG GCCCTGGAGA CCGGCGGAGG GCCAGGACCG CATGAATGC 1020
 CAAAGTGGCG ACTGCAATGG GCACTCAGAG ACATGTCACT TTGACCCCG TGTGTTTGGC 1080
 GCCAGCCAGG GGGCATATGG AGGTGTGTGT GACAATTGCC GGGACCACAC CGAAGGCAAG 1140
 AACTGTGAGC GGTGTGAGCT GCACTATTTT CGGAACCGCG GCCCGGAGG CTTCATTGAG 1200
 65 GAGACCTGCA TCTCTGCGA GTGTGATCCG GATGGGGCAG TGCCAGGGGC TCCCTGTGAC 1260
 CCAGTGACCG GCGAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320
 AAGCGGGGCT TCACTGGACT CACTACGCC AACCGCAGG GCTGCCACCG CTGTGACTGC 1380
 AACATCTGCT GGTCCCGGAG GGACATGCC TGTGACGAGG AGAGTGGGCG CTGCCTTTGT 1440
 CTGCCCAACG TGTGTGGTCC CAAATGTGAC CAGTGTGCTC CTACCACTG GAAGCTGGCC 1500
 70 AGTGGCCAGG GCTGTGAACC GTGTGCTGTC GACCCGCACA ACTCCCTCA GCCCACAGTG 1560
 CAACCACTTC ACAGGCGACT GCCCTGTGCG GAAGGCTTTG GTGGCTGTAT GTGCAGCGCT 1620
 GCAGGCATCC GCGAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAGG ATGCCGAGCC 1680
 TGTGACTGTG ATTTCCGGGG AACAGAGGGC CCGGGCTGCG ACAAGGCATC AGGCCCTGC 1740
 75 CTCTGCGGCC CTGGCTTGAC CGGGCCCGCG TGTGACCACT GCCAGCGAGG CTACTGCAAT 1800
 CGCTACCGCG TGTGCGTGGC CTGCCACCT TGCTTCCAGA CCTATGATGC GGACCTCCGG 1860
 GAGCAGGGCC TGGCTTTGG TAGACTCCGC AATGCCACCG CCAGCTGTGT GTCCAGGCTCT 1920
 GGGCTGGAGG ACCGTGGCCT GGCTCCCGG ATCTAGATG CAAAGAGTAA GATTGAGCAG 1980
 ATCCGAGCAG TTCTCAGCAG CCCCGCAGTC ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT 2040
 80 GCCATCTCTT CCTCAGGGG AACTCTCCAG GGCTTGACG TGGATCTGCC CCTGGAGGAG 2100
 GAGACGTTGT CCTTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGTCTCTCTT 2160
 ACTATGTATC AGAGGAGAGG GGAGCAGTTT GAAAAAATAA GCAGTGTCTG TCTCTCAGGA 2220
 GCCTTCCGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGTCTA GCAGGTCTCC 2280
 GACAGCTCGC GCCTTTTGA CCAGCTCAGG GACAGCCGGA GAGAGGCAGA GAGGCTGGTG 2340
 85 CGGAGGGGAG GAGGAGGAGG AGGCACCGCG AGCCCCAAGC TTGTGGCCCT GAGGCTGGAG 2400
 ATGCTCTGCT TGCTGACCT GACACCCACC TTCAACAGC TCTGTGGCAA CTCCAGGCAG 2460
 ATGGCTTGCA CCCCATATC ATGCCCTGGT GAGCTATGTC CCCAAGCAA TGGCAGACC 2520
 TGTGGCTCCC GCTGCAGGGG TGTCTTCCC AGGCGCGGTG GGGCTTCTT GATGCGGGG 2580
 CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCCAGCTCC AGCGGACCAG GCAGATGATT 2640

5
10
15
20
25

AGGGCAGCCG AGGAATCTGC CTCACAGATT CAATCCAGTG COCAGCGCTT GGAGACCCAG 2700
GTGAGCGCCA GCCGCTCCCA GATGGAGGAA GATGTCAGAC GCACACGGCT CCTAATCCAG 2760
CAGGTCCGGG ACTTCTCTAAC AGACCCCGAC ACTGATGCAG CCACTATCCA GGAGGTGAGC 2820
GAGGCGGTGC TGGCCCTGTG GCTGCCACCA GACTCAGCTA CTGTTCTGCA GAAGATGAAT 2880
GAGATCCAGG CCATTGTCAGC CAGGCTCCCC AACGTGGACT TGGTGTGTC CCAGACCAAG 2940
CAGGACATTG CGCGTGCCCC CCGGTTGCAG GCTGAGGCTG AGGAAGCCAG GAGCCGAGCC 3000
CATCAGATGG AGGCGCAGGT GGAAGATGTG GTTGGGAACC TCGCGCAGGG GACAGTGGCA 3060
CTGCAGGAAG CTCAGGACAC CATGCAAGGC ACCAGCCGCT CCCTTCGGCT TATCCAGGAC 3120
AGGGTTGCTG AGGTTTACGA GGTACTGCGG CCAGCAGAAA AGCTGGTGAC AAGCATGACC 3180
AAGCAGCTGG GTGACTTCTG GACACGGATG GAGGAGCTCC GCCACCAAGC CCGGCAGCAG 3240
GGGGCAGAGG CAGTCCAGGC CCAGCAGCTT GCGGAAGGTG CCAGCGAGCA GGCATTGAGT 3300
GCCCAAGAGG GATTGTAGAG AATAAAACAA AAGTATGCTG AGTTGAAGGA CCGGTTGGGT 3360
CAGAGTTCCA TGCTGGGTGA GCAGGGTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG 3420
GAGCTGTTT GGCAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480
CTGCGGGGCA GCCAGGCCAT CATGCTGCGC TCGCGGAGCC TGACAGGACT GGAGAAGCGT 3540
GTGGAGCAGA TCCGTGACCA CATCAATGGG CGCGTGCTCT ACTATGCCAC CTGCAAGTGA 3600
TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCGCGC TTGCTTTTGT GTTGGGGGCA 3660
GATTGGGTGT GAATGCTTTC CATCTCCAGG AGACTTTTAT GCAGCCTAAA GTACAGCCTG 3720
GACCAACCCCT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAAAT 3780
GGGACAGTTA CACTTGACAG ACAAGATGG TGGAGATTGG CATGCCATTG AAACCTAAGAG 3840
CTCTCAAGTC AAGGAAGCTG GGCTGGGCGG TATCCCCCGC CTTTAGTTCT CCACTGGGGA 3900
GGAATCCTCG ACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAAATGAA AAGCCAAATA 3960
AAAATCTTTG G

Seq ID NO: 521 Protein sequence
Protein Accession #: NP_000219.1

30
35
40
45
50

1 11 21 31 41 51
MRPFFLLCPA LPGLLHAQQA CSRGACYPPV GDLLVGRTRP LRASSTCGLT KPETYCTQYG 60
EWQMKCKCD SRQPHNYSH RVENVASSSG PMRWQSQND VNPVSLQLDL DRRFPLQEVN 120
MEFGPMFAG MLIERSSDFG KTWVYQYLA ADCTSTFPRV RQGRPOSWQD VRCQSLPQRP 180
NARLNGKQVQ LNLMDLVSGI PATQSQKIQE VGEITNLRVN FTRLAPVPQR GYHPPSAYYA 240
VSQRLQGSC FCHGHADRC PKPGASAGPS TAVQVHDVCV CQHNTAGPNC ERCAFFYNNR 300
PWRFAEQQDA HECQRCDGNG HSETCHFDPA VFAASQGAAY GVCNCRDHT EGKNCERQQL 360
HYFNRRRPGA SIQETCISCE CDPDGAVPGA PCDPVTGQCV CKEHVQGERC DLCKPGFTGL 420
TYANPQCHSR CDCNLLGSRP DMPDDEESGR CLCLPNVVG P KCDQCAPYHW KLASGQCEP 480
CACPENSPQ PTVQPVHRAV PCREGFGLM CSAAIRQCP DRTYGDVATG CRACDCDFRG 540
TEGPGCDKAS GRCLCRPGLT GPRCDQCQRG YCNRYPCVCA CHPCFQTYDA DLREQALRFG 600
RLRNATASLW SGPLCEDRGL ASRILDAKSK IEQIRAVLSS PAVTEQEVQA VASAILSLRR 660
TLQGLQLDLP LEBETLSLPR DLESILDRSFN GLLTMYQRKR EQFEKISSAD PSGAFRLMT 720
AYEQSAQAQ QVSDSSRLLD QLRDSRREAE RLVRQAGGGG GTGSPKLVAL RLEMSLPLD 780
TFPTFNKLCGN SRQMACTPIS CPGLCPQDN GTACGSRERG VLPFRAGGAFI MAQVQAEQLR 840
GFNAQLQRTR QMIRAAEESA SIQISSAQLR ETQVSASRSQ MEEDVRRTRL LIQVVRDFLT 900
DPDIDAATIQ QMSEAVLALW LPTDSATVLQ KMWELQIAIA RLPNVDLVLS QTKQDIARAR 960
RLQAEAEER SMRAHVEQGV EDVVGNLROG TVALQEAQDT MQGTSRSLRL IQDRVAEVQQ 1020
VLRFPAERLVT MTKQLGDFW TRMEELRHQA RQOGAEVQA QQLAEGASEQ ALSAQBGFER 1080
IKQRYAELKD RLQSSMLGE QQARIQSVKT BAEELFGETM ENMDRMKDME LELLRGSQAI 1140
MLRSADLTGL EKRVEQIRDH INGRVLYYAT CK

Seq ID NO: 522 DNA sequence
Nucleic Acid Accession #: NM_001944.1
Coding sequence: 84..3083

55
60
65
70
75
80
85

1 11 21 31 41 51
TTTCTTAGA CATTAACATGC AGACGGCTGG CAGGATAGAA GCAGCGGCTC ACTTGGACTT 60
TTTACCAGG GAAATCAGAG ACAATGATGG GGCTCTTCCC CAGAACTACA GGGGCTCTGG 120
CCATCTTCGT GGTGGTCATA TTGGTTTCATG GAGAATTGGG AATAGAGACT AAAGGTCAAT 180
ATGATGAAGA AGAGATGACT ATGCAACAAG CTAAAGAAGG GCAAAAACGT GAATGGGTGA 240
AATTTGCCAA ACCCTGCAGA GAAGGAGAAG ATAACTCAA AAGAAACCCA ATTGCCAAGA 300
TTACTTCAGA TTACCAAGCA ACCCAGAAAA TCACCTACCG AATCTCTGGA GTGGGAATCG 360
ATCAGCCGCC TTTTGAATC TTTGTTGTTG ACAAAAACAC TGGAGATATT AACATAACAG 420
CTATAGTCGA CCGGGAGGAA ACTCCAAGCT TCCTGATCAC ATGTCGGGCT CTAAATGCCC 480
AAGGACTAGA TGTAGAGAAA CCACTTATAC TAACGGTTAA AATTTTGGAT ATTAATGATA 540
ATCCTCCAGT ATTTTCAACA CAAATTTTCA TGGGTGAAT TGAAGAAAT AGTGCTCAA 600
ACTCAGTGGT GATGATACCT AATGCCACAG ATGCAGATGA ACCAAACCA CTGAATTCTA 660
AAATTGCCTT CAAATTTGTC TCTCAGGAAC CAGCAGGCAC ACCCATGTTT CTCTAAGCA 720
GAAACACTGG GGAAGTCCGT ACTTTGACCA ATTCTCTTGA CCGAGAGCAA GCTAGCAGCT 780
ATCGTCTGGT TGTGAGTGGT GCAGACAAAG ATGGAGAAGG ACTATCAACT CAATGTGAAT 840
GTAATATTAA AGTGAAGAT GTCAACGATA ACTTCCCAAT GTTTAGAGAC TCTCAGTATT 900
CAGCAOAGT TGAAGAAAT ATTTTAAOTT CTGAATTACT TCGATTTCAG GTAACAGATT 960
TGGATGAAGA GTACACAGAT AATTGGCTTG CAGTATATT CTTTACCTCT GGGAAATGAAG 1020
GAAATTGGTT TGAATACAA ACTGATCCTA GAACTAATGA AGGCATCCTG AAAGTGGTGA 1080
AGGCTCTAGA TTATGAACAA CTACAAAGCG TGAACCTTAG TATTGCTGTC AAAAACAAAG 1140
CTGAATTTCA CCAATCAGTT ATCTCTCGAT ACCGAGTTCA GTCAACCCCA GTCACAATTC 1200
AGGTAATAAA TGTAAGAGAA GGAATTGCAT TCCGTCTGCG TTCCAAGACA TTTACTGTGC 1260
AAAAAGGCAT AAGTAGCAAA AAATTGGTGG ATTATATCCT GGAACATAT CAAGCCATCG 1320
ATGAGGACAC TAACAAGCT GCCTCAATG TCAAAATATG CATGGGACGT AACGATGGTG 1380
GATACCTAAT GATTGATTCA AAAACTGCTG AAATCAAAAT TGTCAAAAAT ATGAACCGAG 1440
ATTCTACTTT CATAGTTAAC AAAACATCA CAGCTGAGGT TCTGGCCATA GATGAATACA 1500
CGGGTAAAA TTCTACAGGC ACGTATATG TTAGAGTACC CGATTTCAAT GACAATTGTC 1560
CAACAGCTGT CCTCGAAAA GATGCAGTTT GCAGTTCTTC ACCTTCGGTG GTTGTCTCCG 1620
CTAGAACACT GAATAATAGA TACACTGGCC CTTATACATT TGCACTGGAA GATCAACCTG 1680
TAAAGTTGCC TCCGCTATGG AGTATCAAA CCTCAATGC TACCTCGGCG CTCTCTCAGG 1740
CCAGGAACA GATACCTCT GGAGTATACC ACATCTCCCT GGTACTTACA GACAGTCAGA 1800
ACAATCGGTG TGAGATGCCA CGCAGCTTGA CACTGGAAGT CTGTCACTGT GACAACAGGG 1860

	GCATCTGTGG	AACCTCTTAC	CCAACCACAA	GCCCTGGGAC	CAGGTATGGC	AGGCGGCACT	1920
	CAGGGAGGCT	GGGGCCCTGCC	GCCATCGGCC	TGCTGCTCCT	TGGTCTCCTG	CTGCTGCTGT	1980
	TGGCCCCCTC	TCTGCTGTGG	ACCTGTGACT	GTGGGGCAGG	TTCTACTGGG	GGAGTGACAG	2040
5	GTGGTTTAT	CCAGTTCTCT	GATGGCTCAG	AAGGAACAAT	TCATCAGTGG	GGAAITGAAG	2100
	GAGCCCATCC	TGAAGACAAG	GAAATCACAA	ATATTGTGT	GCCTCCTGTA	ACAGCCAATG	2160
	GAGCCGATT	CATGGAAAGT	TCTGAAGTTT	GTACAAATAC	GTATGCCAGA	GGCAGAGGG	2220
	TGGAAAGCAC	TTCAGGAATG	GAAATGACCA	CTAAGCTTGG	AGCAGCCACT	GAATCTGGAG	2280
	GTGCTGCAGG	CTTTGCACAA	GGGACAGTGT	CAGGAGCTGC	TTGAGGATTC	GGAGCAGCCA	2340
10	CTGGAGTTGG	CATCTGTTC	TCAGGCGAGT	CTGGAACCAT	GAGAACAAAG	CATTCCACTG	2400
	GAGGAACCAA	TAAGGACTAC	GCTGATGGGG	CGATAAGCAT	GAATTTTCTG	GACTCCTACT	2460
	TTTCTCAGAA	AGCATTGTGC	TGTGCGGAGG	AAGACGATGG	CCAGGAAGCA	AATGACTGCT	2520
	TGTTGATCTA	TGATAATGAA	GGCGCAGATG	CCACTGGTTC	TCCTGTGGGC	TCOGTGGGTT	2580
	GTTGCACTTT	TATTGCTGAT	GACCTGGATG	ACAGCTTCTT	GGACTCACTT	GGACCCAAAT	2640
15	TTAAAAAAT	TGCAGAGATA	AGCCTTGGTG	TTGATGGTGA	AGGCAAGAA	GTTCAAGCCAC	2700
	CCCTAAAGA	CAGCGGTTAT	GGGATTGAAT	CCTGTGGCCA	TCCCATAGAA	GTCAGCAGAG	2760
	CAGGATTGTG	TAAGTGCCAG	ACTTTGTGAG	GAAGTCAAGG	AGCTTCTGCT	TTGTCCGCTT	2820
	CTGGGTCTGT	CCAGCCAGCT	GTTTCCATCC	CTGACCTCTT	GCAGCATGGT	AACTATTTAG	2880
	TAAAGGAGAC	TACTCGGCTT	TCTGTTCCCT	TCGTGCAACC	TTCACTGCA	GGCTTTGATC	2940
20	CACCTCTCAC	ACAAAAATGT	ATAGTGACAG	AAAGGCTGAT	CTGTCCCAT	TCCAGTGTTC	3000
	CTGGCAACTC	AGCTGGGCGA	ACGCAGCTAC	GAGGCTCACA	TACTATGCTC	TGTACAGAGG	3060
	ATCCTTGCTC	CCGTCTAATA	TGACCAGAAT	GAGCTGGAAT	ACCACACTGA	CCAAATCTGG	3120
	ATCTTTGGAC	TAAAGTATTC	AAAAATAGCAT	AGCAAAGCTC	ACTGTATTGG	GCTAATAATT	3180
	TGGCACTTAT	TAGCTTCTCT	CATAAACTGA	TCACGATTAT	AAATTAAATG	TTTGGGTTCA	3240
25	TACCCCAAAA	GCAATATGTT	GTCACCTCTA	ATTCTCAAGT	ACTATTCAAA	TTGTAGTAAA	3300
	TCTTAAAGTT	TTTCAAAACC	CTAAAAATCAT	ATTCGC			

Seq ID NO: 523 Protein sequence
Protein Accession #: NP_001935.1

30	1	11	21	31	41	51	
	MMGLFPRTTG	ALAFVUVVIL	VHGLRIETK	QYDEEEMTM	QAKRRQKRE	WVKFAPCRE	60
	GEDNSKRNP	AKITSDYQAT	QKITYRISGV	GIDQPPFGIF	VVDKNTGDI	ITAIVREBT	120
	PSFLITCRAL	NAQLDVEKP	LILTVKILDI	NDNPVVFSSQ	IFMGEIEENS	ASNSLVMIIN	180
35	ATDADEPNFL	NSKTAPIKIV	QEPAGTMMFL	LSRNTGEVRT	LTNSLDREQA	SSYRLVVSGA	240
	DKDGEGLSTQ	CECNKIKVDV	NDNPFMFRTS	QYSARIEENI	LSSELLRPQV	TDLDSEYTDN	300
	WLAVVYFTSG	NEGNWFEIQT	DPRINBGLIK	VVKALDYBOL	QSVKLSIAVK	NKAEFHQSVI	360
	SRVRVQSTPV	TIQVINVREG	IAFRPASKTF	TVQKGISSKK	LVYILGTYQ	AIDEDINKAA	420
	SNVKYVMGRN	DGGYLMIDSK	TAEIKFVKNM	NRDSTFIVNK	TITAEVLAI	EYTGKISTGT	480
40	VYVRVPDFND	NCPFAVLEKD	AVCSSSPSVV	VSARTLNNRY	TGFTYFALED	QPVKLPAVNS	540
	ITTLNATSL	LRAGBEQIPP	VYHISLVLT	SONNRCMPR	SLTLEVCCQD	NRGICSTSY	600
	TTSPGTTRYG	PHSGRLGPAA	IGLLLLGLLL	LLLAPLLLL	CDGAGSTGG	VTGGFIPVPD	660
	GSEGTIHQWG	IEGAHPEDKE	ITNICVPPVT	ANGADFMES	EVCTNTYARG	TAVEGTSME	720
	MTTLKLAATE	SGGAAGFATG	TVSGAASGFG	AATGVGICSS	GGSGTMRTRH	STGGTNKDYA	780
45	DGAISMNPLD	SYFSQKAFAC	AEEDDQGEAN	DCLLIYDNB	ADATGSPVGS	VGCCSFIADD	840
	LDDSFILDSL	PKFKKLAEIS	LGVDGEGKEV	QPPSKDSGYG	IESCGHPIEV	QQTGFVKCQT	900
	LSSGQASAL	SASGSVPQAV	SIPDLQHCN	YLVTTYSAS	GSLVQPSFAG	FDPLLTONVI	960
	VTERVICPIS	SVPGNLGAPT	QLRGSHTMLC	TEDPCSRLL			

Seq ID NO: 524 DNA sequence
Nucleic Acid Accession #: XM_058069.2
Coding sequence: 1..1413

55	1	11	21	31	41	51	
	ATGAAGITTC	TTCTAATACT	GCTCCTGCAG	GCCACTGCTT	CTGGAGCTCT	TCCCTGAAC	60
	AGCTCTACAA	GCCTGGAAAA	AAATAATGTG	CTATTGGTGG	AAAGATACCT	AGAAAAATTT	120
	TATGGCCCTG	AGATAAACAA	ACTTCCAGTG	ACAAAAATGA	AAATAGTGG	AAACTTAATG	180
	AAGGAAAAAA	TCCAGAAAT	GCAGCACTTC	TTGGGTCTGA	AAGTGACCGG	GCAACTGGAC	240
60	ACATCTACCC	TGGAGATGAT	GCAAGCACTT	CGATGTGGAG	TCCCGATGT	CCATCATTTT	300
	AGGGAAATGC	CAGGGGGGCC	CGTATGGAGG	AAACATTATA	TCACCTACAG	AATCAATAAT	360
	TACACACCTG	ACATGAACCG	TGAGGATGTT	GACTACGCAA	TCCGGAAGC	TTTCCAAGTA	420
	TGGAGTAATG	TTACCCCTCT	GAAATTCAGC	AAGATTAAAC	CAGGCATGGC	TGACATTTTG	480
	GTGGTTTTTG	CCCGTGGAGC	TCATGGAGAC	TTCCATGCTT	TTGATGGCAA	AGGTGGAATC	540
65	CTAGCCCATG	CTTTTGGAGC	TGGATCTGGC	ATTGGAGGGG	ATGCACATTT	CGATGAGGAC	600
	GAATTCCTGA	CTACACATTC	AGGAGGCACA	AACTTGTTC	TCAGTCTGT	TCACGAGATT	660
	GGCCATTCCT	TAGGCTCTTG	CCATCTTAGT	GATCCAAAGG	CGTAATGTT	CCCCACCTAC	720
	AAATATGTTG	ACATCAACAC	ATTTGCGCTC	TCTGCTGATG	ACATACGTGG	CATTCACTCC	780
	CTGTATGGAG	ACCCAAAGAG	GAACCAACGC	TTGCCAAATC	CTGACAAATC	AGAACCAAGT	840
70	CTCTGTGACC	CCAATTTGAG	TTTTGATGCT	GTCACTACCG	TGGGAAATAA	GATCTTTTTC	900
	TTCAAAGACA	GTTTCTCTGT	GCTGAAGGTT	TCTGAGAGAC	CAAAGACCAG	TGTTAATTTA	960
	ATTTCTTCCT	TATGGCCAAC	CTTGCCATCT	GGCAITGAAG	CTGCTTATGA	AATTGAAGCC	1020
	AGAAATCAAG	TTTTTCTTTT	TAAAGATGAC	AAATACTGGT	TAATTAGCAA	TTTAAGACCA	1080
	GAGCCAAAT	ATCCCAAGAG	CATACATCT	TTTGGTTTTT	CTAATCTTGT	GAAAAAATTT	1140
75	GATGCAGCTG	TTTTTAACCC	ACGTTTTTAT	AGGACCTACT	TCCTTGTAGA	TAACCAATAT	1200
	TGGAGGTATG	ATGAAGGAG	ACAGATGATG	GACCTGTGTT	ATCCCAACT	GATTACCAAG	1260
	AACTTCCAAG	GAATCGGGCC	TAAAATTGAT	GCACTCTTCT	ACTCTAAAAA	CAATACTAC	1320
	TATTTCTTCC	AAGGATCTAA	CCAATTTGAA	TATGACTTCC	TACTCCAACG	TATCACCAAA	1380
80	ACACTGAAAA	GCAATAGCTG	GTTTGGTTGT	TGA			

Seq ID NO: 525 Protein sequence
Protein Accession #: P39900

85	1	11	21	31	41	51	
	MKFLILLILQ	ATASQALPLN	SSTSLEKNV	LFGERYLEKF	YGLEINKLEV	TKMKYSGNLM	60
	KEKIQEMQHF	LGLKVTGQLD	TSTLEMMHAP	RCGVFVDFHF	REMPGGPVWR	KHYITYIRIN	120

YTPDMNREUV DYAIRKAPQV WSNVTPLKFS KINTGMADIL VVFARGAHGD FHAFDGKGGI 180
 LAHAFPGSGG IGGDAHPDED EFWTHSGGT NLFLTAVHEI GHSGLGLGHS DPKAVMFPTY 240
 KYVDINTFRL SADDIRGIQS LYGDPKENQR LPNPDNSEPA LCDPNLSFDA VTTVGNKIFP 300
 FKDRPFWLKV SERPKTSVNL ISSLWPTLPS GIEAAVEIEA RNQVFLPKDD KYWLISNLRP 360
 5 EPNYPKSIHS FGPFNPVKKI DAAVFNPRFY RTYFFVDNQY WRYDERRQMM DPGYFKLITK 420
 NFGGIGPKID AVFYSKNKY YFFQGSNQFE YDFLLQIRTK TLKSNSWFGC

Seq ID NO: 526 DNA sequence

Nucleic Acid Accession #: NM_024423.1

Coding sequence: 64..2590

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

1	11	21	31	41	51	
GGCAGGTC	GCTCTCGGCA	CCCTCCCGGC	GCCCGCGTTC	TCCTGGCCCT	GCCCGGCATC	60
COGATGGCG	CCGCTGGGCC	CCGCGCGCTCC	GTGCGCGGAG	CCGCTGCGCT	GCATCTGCTG	120
CTGACCCCTG	TGATCTTCAG	TCGTGATGGT	GAAGCGCTGCA	AAAAGGTGAT	ACTTAAATGTA	180
CCCTCTAAAC	TAGAGGCAGA	CAAAATAAAT	GGCAGAGTTA	ATTGGGAAGA	GTGCTTCAGG	240
TCCTGACACC	TCATCCGGTC	AAGTGATCCT	GATTTCAGAG	TTCTAAATGA	TGGGTTCAGTG	300
TACACAGCCA	GGGCTGTTGC	GCTGTCTGAT	AAGAAAAGAT	CATTATACCAT	ATGGCTTTCT	360
GACAAAAGGA	AACAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAAGAAGGTA	420
TCGAAGACAA	GACACACTAG	AGAAACCTGT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCACCT	480
ATTCCCTGGCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTCC	CATTGTTTCT	TCAACAAGTT	540
GAATCTGATG	CAGCAGAGAA	CTATACCTGC	TTCTACTCAA	TAAGTGAACG	TGGAGTTGAT	600
AAAGAACCCT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAAATCTATT	TTGCACTCGG	660
CCTGTGATGC	GTGAAGAATA	TGATGTTTTT	GATTTGATTG	CTTATGCGTC	AACATGCAGAT	720
GGATATTCCG	CAGATCTGCC	CCTCCCACTA	CCCATCAGGG	TAGAGGATGA	AAATGACAAC	780
CACCCGTGTT	TCACGAAGAG	AATTTATAAT	TTTGAAGTTT	TGGAAAGTAG	TAGACCTGGT	840
ACTACAGTGG	GGGTGTTTGG	TGCCACAGAC	AGAGATGAAC	CGGACACAAT	GCATACGCGC	900
CTGAAATACA	GCATTTTGCA	GCAGACACCA	AGGTCACCTG	GGCTCTTTTC	TGTGCATCCC	960
AGCACAGGGC	TATCTACCCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
TCATTGATAA	TGAAGTACAC	AGACATGGAT	GGCCAGTTTT	TTGATTGATG	AGGCACATCA	1080
ACTTGTATCA	TACAGTAAAC	AGATTCAAAT	GATAATGCAC	CCACTTTCAG	ACAAAATGCT	1140
TATGAAGCAT	TTGTAGAGGA	AAATGCATTC	AATGTGGAAA	TCCTACGAAT	ACCTATAGAA	1200
GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTTT	AAAGGGAAAT	1260
GAAAATGGAC	ATTTCAAAAT	CAGCACAGAC	AAAGAACTA	ATGAAGGTGT	TCTTTCTGTT	1320
GTAAAGCCAC	TGAATATAGA	AGAAAACCGT	CAAGTGAACC	TGGAATTTGG	AGTAAACAAT	1380
GAAGCGCAT	TTCTAGAGAA	TATCCACAGA	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
GTTCATGTGA	GGGATCTGGA	TGAGGGGCGT	GAATGCATCT	CTGCAGCCCA	ATATGTGCGG	1500
ATTAAGGAAA	ACTTAGCAGT	GGGGTCAAGG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCAAT	ATCCTAAAGG	TTGGATCACC	1620
ATTGATGAAA	TTTCAGGGTG	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTGAAACT	1680
CCCAAAATGT	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAAATGATA	ATCCACCAGA	AATACTTCAA	1800
GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAATTTGC	CCAATACTTC	TCCAGAAATC	1920
AGTAGACTGT	GGAGCCTCAC	CAAGTTAAT	GATACAGCTG	CCGCTCTTTC	ATATCAGAAA	1980
AATGCTGGAT	TTCAAGAAAT	TACCAATCCT	ATTACTGTAA	AAGACAGGGC	CGGCGCAAGT	2040
GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACCTCA	GTGCTGTGGG	2100
ACTTCAAGGA	GTACAGAGGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
ATAGCACTGC	TCTTTCTGCT	ATTGCTAACT	TAGTATGTGT	GAGTTTTTGG	TGCAACTAAA	2220
GGGAAACGTT	TTCTCGAAGA	TTTAGCACAG	CAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
CCTGGAGACG	ATAGAGTGTG	CTCTGCCAAT	GGATTATGCA	CCCAAACTAC	CAACAACCTC	2340
AGCCAAGGTT	TTTGTGGTAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAACCAAT	2400
GAAATAGTGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
ACCCTGGACT	CCTGCGGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACCTACTCG	2520
GAGTGGCACA	GTTTACTCA	ACCCCGTCTC	GGTGAAGAA	CCATTAGAGG	ACACACTGGT	2580
TAAAAATTA	ACATAAAGA	AATTCATCG	ATGTAATCAG	AATGAAGACC	GCATGCCATC	2640
CCAGATTAT	TGCTCACTT	ATAACTATGA	GGGAAGAGGA	TCTCCAGCTG	GTTCGTGGG	2700
CTGCTCGAGT	GAAAGCAGG	AAGAGATGG	CCTTGACTTT	TTAAATAAAT	TGGAACCCAA	2760
ATTATTATCA	TTAGCAGAA	CATGCACAAA	GAGATAATGC	CACAGTGCTA	CAATTAGGTC	2820
TTTGTGAGAC	ATTCTGGAGG	TTTCAAAAA	TAATATTGTA	AAGTTCAATT	TCAACATGTA	2880
TGTATATGAT	GATTTTCTT	TCAATTTTGA	ATTATGCTAC	TCACCAATTT	ATATTTTTAA	2940
AGCCAGTTGT	TGCTTATCTT	TTCCAAAAAG	TGAAAAATGT	TAAAAACAGC	AACCTGGTAA	3000
TCTCAAACTC	CAGCACTGGA	ATTAAAGTCT	CTAAAGCATC	TGCTCTTTTT	TTTTTTTACG	3060
GATATTTTAG	TAATAAATAT	GCTGGATAAA	TATTAGTCCA	ACAATAGCTA	AGTTATGCTA	3120
ATATCATCAT	ATTATGTATT	CACTTTAAGT	GATAGTTTAA	AAAATAAACA	AGAAATATTG	3180
AGTATCATA	TGTGAAGAAA	GTTTGGAAA	AGAAACAATG	AAGACTGAAT	TAAATTAATA	3240
ATGTTGCAGC	TCATAAAGAA	TGGGACTCA	CCCTACTGCG	ACTACCAAT	TCATTGACT	3300
TTGGAGGCAA	AATGTGTTGA	AGTGCCCTAT	GAAGTAGCAA	TTTTCTATAG	GAATATAGTT	3360
GGAAATAAAT	GTGTGTGTGT	ATATTATTAT	TAATCAATGC	AATATTTAAA	ATGAAATGAG	3420
AACAAAGAGG	AAATGGTAA	AAACTTGAAA	TGAGGCTGGG	GTATAGTTTG	TCCTACAATA	3480
GAAAAAAGAG	AGAGCTTCCT	AGGCTGGGCG	TCTTAAATGC	TGCATTATAA	CTGAGTCTAT	3540
GAGGAATAG	TTCTGTCCA	ATTGTGTAA	TTTGTTTAAA	ATTGTAAATA	AATTAAACTT	3600
TTCTGTTTTC	TGTGGGAAGG	AAATAGGGAA	TCCAATGGAA	CAGTAGCTTT	GCTTTGCAGT	3660
CTGTTTCAAG	ATTCTGCAAT	CCACAAGTTA	GTAGCAAACT	GGGGAATACT	CGCTGCAGCT	3720
GGGGTTCCCT	GCTTTTGTGT	AGCAAGGGTC	CAGAGATGAG	GTGTTTTTTT	CGGGGAGCTA	3780
ATAACAAAA	CATTTTAAAA	CTTACCTTTA	CTGAAGTTAA	ATCCTCTATT	GCTGTTTCTA	3840
TTCTCTCTTA	TAGTAGCCAA	CATCTTTTAA	ATTAGATACC	AAATAACCAT	GTCTCTCTAG	3900
AGTTTAGAGG	CTAGAGGGAG	CTGAGGGGAG	GATCTTACTG	AAAGCACCTT	GGGGAGATTG	3960
ATTGTCTTAA	TAACCTAAGC	CCACAAACTT	GACACCTGAT	CAGGTCTGGG	AGCTACAAAA	4020
TTTCATTTTT	CTCCTCACTG	CCCTTCTTCT	GAGTGGCATT	GGCCTGAATC	AAGGAAAGCC	4080
AGGCCCTTGT	GGCCCCCTTC	TTTGGGCTTT	CTGCTAAAGC	AACACCTCCA	GCAGAGATTG	4140
CCTTAAGTGA	CTCCAGGTTT	TCCACCATCC	TTAGCGTGA	ATTAAATTTT	AATCAGTTTG	4200
CTTCTCTCCG	AGAAATTTTA	AAATAATAGA	AGAAATAGAA	ATTTTGAATG	TATAAAGGAA	4260
AAAGATCAAG	TTGTCTTTTT	AGAACAGAGG	GAACCTTGGG	AGAAAGCAGC	CCAAGTAGGT	4320
TATTTGTACA	GTACAGGGGC	AACAGGAAGA	TGCAGGCCCT	CAAGGGCAAG	GAGAGGCCAC	4380
AAGGAATATG	GGTGGGAGTA	AAAGCAACAT	CGTCTGCTTC	ATACTTTTTTC	CTAGGCTTGG	4440

5	CACCTGCTTT	TCCTTTCTCA	GGCCAATGGC	AACCTGCCATT	TGAGTCCGGT	GAGGGATCAG	4500
	CCAACCTCTT	CTCTATGGCT	CACCTTATTT	GGAGTGAGAA	ATCAAGGAGA	CAGAGCTGAC	4560
	TGCATGATGA	GTCTGAAGGC	ATTTGCAGGA	TGAGCCTGAA	CTGGTTGTGC	AGAACAAACA	4620
	AGGCATTCAAT	GGGAATTGTT	GTATTCTTTC	TGCAGCCCTC	CTTCTGGGCA	CTAAGAAGGT	4680
	CTATGAATTA	AATGCCATATC	TAAAAATTCTG	ATTTATTCTCT	ACATTTTCTG	TTTTCTAATT	4740
	TGACCCCTAAA	ATCTATGTGT	TTTAGACTTA	GACTTTTTAT	TGCCCCCCCC	CCCTTTTTTT	4800
	TTAGAGCGGA	GTCTCGCTCT	GACGCACAGG	CTGGAGTGCA	GTGGCTCCGA	TCTCTGTCTA	4860
	CTGAAGGCTC	CGCTCCCGGC	GTTTCATGCCA	TTCTCTGCCC	TCAGCCTCCT	GAGTAGCTGG	4920
10	GACTAGAGGC	GCCACCAACC	ACGCCCGGCT	AATTTTTTGT	ATTTTATAA	GAGACGGGGT	4980
	TTCACTGTGT	TAGCCAGGAT	GGTCTCGATC	TCTTGACCTC	GTGATCCGCC	TGCCCTGGCC	5040
	TCCCAAAGTG	CTGGGATTAC	AGGCATGACC	CACCGCTCCC	GGCCTTGTTT	TCCGTTTAAA	5100
	GTCGTCTTCT	TTTAATGTAA	TCATTTTGAA	CATGTGTGAA	AGTTGATCAT	ACGAATTGGA	5160
	TCAATCTTGA	AATACTCAAG	CAAAAGACAG	TCGAGAAGCC	AGGGGGAGAA	AGAACTCAGG	5220
	GCACAAATA	TTGGTCTGAG	AATGGAATTC	TCTGTAAGCC	TAGTTGCTGA	AATTTCTGTC	5280
15	TGTAACCAGA	AGCCAGTTTT	ATCTAACGGC	TACTGAAACA	CCCACTGTGT	TTTGCTCACT	5340
	CCCCTCACC	GATCAAAACC	TGCTACCTCC	CCAAGACTTT	ACTAGTCCCG	ATAAACTTTC	5400
	TCAAAGAGCA	ACCAGTATCA	CTTCCCTGTT	TATAAAACCT	CTAACCATCT	CTTTGTTCTT	5460
	TGAACATGCT	GAACAACACC	TGGTCTGCAT	GTATGCCCGA	ATTGTAATT	CTTTCTCTCT	5520
	AAATGAAAA	TTAATTTTAG	GGATTCAATT	CTATATTTTC	ACATATGTAG	TATTATTATT	5580
20	TCCTTATATG	TGTAAGGTGA	AATTTATGGT	ATTTGAGTGT	GCAAGAAAA	ATATTTTTAA	5640
	AGCTTTCATT	TTTCCCCCAG	TGAATGATTT	AGAAATTTTT	ATGTAAATAT	ACAGAATGTT	5700
	TTTTCTTACT	TTTATAAGGA	AGCAGCTGTC	TAAATGCAG	TGGGTTTGT	TTTGCAATGT	5760
	TTTAAACAGA	GTTTGTATAT	TGCTATTAAA	AGAAGTTACT	TTGCTTTTAA	AGAACTTGG	5820
	CTGCTTAAAA	TAAGCAAAA	TTGGATGCAT	AAAGTAATAT	TTACAGATGT	GGGGAGATGT	5880
25	AATAAACA	TATTAACTTG	GCTGCTTAAA	ATAAGCAAAA	ATTGGATGCA	TAAAGTAATA	5940
	TTTACAGATG	TGGGGAGATG	TAAATAAACA	ATATTAACTT	GGTTTCTTGT	TTTTGTCTGA	6000
	TTTAGAGATT	AAATAATTCT	AAGATGATCA	CTTTGCAAAA	TTATGCTTAT	GGCTGGCATG	6060
	GAATAGAAA	TACTCAATTA	TGCTTTTGT	GTATTAATGG	GGATATTTT	GGACAATGTT	6120
	TCATTATCAA	ATTGTCGACA	TCATTAAATAT	ATATTGTAAT	GTGGGAAGA	GATCACTATT	6180
30	TTGAAGCACA	GCTTTACAGA	TGAGTATCTA	TGATACATAT	GTATAATAAA	TTTTGATCGG	6240
	GTATTAAAGT	TATTAGAAGG	TGGTTATAAT	TGCAGAGTAT	TCCATGAATA	GTACACTGAC	6300
	ACAGGGGTTT	TACTTTGAGG	ACCAGTGTAG	TCAAGGGAAA	ACATGAGTTA	AAAAGAAAAG	6360
	CAGGCAATAT	TGCAGTCTTG	ATTCTGCCAC	TTACAGGATA	GATAATGCCT	GAACTTTAAT	6420
35	GACAAGATGA	TCCAACCTCA	AAGGTGCTCT	GTGCTTCACA	GTGAATCTTT	TCCCCATGCA	6480
	GGAGTGTGCT	CCCTTACAAA	CGTTAAGACT	GATCATTTCA	AAAATCTATT	AGCTATATCA	6540
	AAAGCCTTAC	ATTTTAAATAT	AGGTTGAACC	AAAAATTTCA	TTCCAGTAAC	TTCTATTGTA	6600
	ACCATTAATT	TTTGTTATGT	CTTCAAGAAT	GTTTATGGA	TTTTTGTGTT	TAATAGTAAA	6660
	ATACCGGATA	CATTTCACGT	GTCCTTCAGT	ATTGATTGG	TTGAATATTG	GGTCATAATG	6720
	GTTGAGAAGC	ATGGACACTA	GAGCCAGAAT	GCTTGGATAT	GAATCCTGGA	TCTGTCACTT	6780
40	ACTTCTGTGT	GACCTTTGAA	AGGCTACTTA	TTTCTCTCT	TAGCTTTCTC	ATTAAAAATCA	6840
	ATGAACAATG	CCAGCCTCAT	GGGGTTGTTG	AATGATTAAA	TTAGTTAATA	TACCTAAAGT	6900
	ACATAGAAC	CTGCTGCAC	ATAGTAAAG	AATTATAAGT	GTGAGGTAGT	TGGTAAATTT	6960
	ATGTAGTTGG	ATATACTACC	GAACAATATC	TAACTCTCTT	TTAGGGAAAT	AAAGTTTGTG	7020
45	CATATATATA	ATCCCGAAAC	ATG				

Seq ID NO: 527 Protein sequence
Protein Accession #: NP_077741.1

50	1	11	21	31	41	51	
	MAAAGPRRSV	RGAVCLHLL	TLVIFSRDGE	ACKKVILNVP	SKLEADKIIG	RVNLEECFRS	60
	ADLRSSDPD	FRVLNDGSV	TARAVALS	KRSFTIWLSD	KRKQTKQEV	VLEHQKKVS	120
	KTRHRTETVL	RRAKRRWAPI	PCSMQENSLG	PPPLFLQOVE	SDAAQNYTVF	YSISGRGVDK	180
	EPLNLFYIER	DTGNLFCRTP	VDREEDVDFD	LIAYASTADG	YSADLPLPLP	IRVEDENDNH	240
55	PVPTFAIYNF	EVLESSRPGT	TVGVVVCATDR	DEPDTMHTRL	KYSILQQTTPR	SPGLFSVHPS	300
	TGVITTVSHY	LDREVDVKYS	LIMKVQDMDG	QFFFLIGTST	CIITVDSND	NAPTFRQNAV	360
	EAFVEENAFN	VEILRIPIED	KDLINTANWR	VNFTILKQNE	NGHPRISTD	ETNEGVLSVV	420
	KPLNYEENRQ	VNLEIGVNN	APFARDIPRV	TALNRALVT	HVRDLDEGPE	CTPAAQYVRI	480
	KENLAVGSKI	NGYKAYDPEN	RNGNGLRYKK	LHDPKGMWIT	DEISGSIITS	KILDREVEPT	540
60	KNELYNITVL	AIDRDRSCT	GTLANVIEDV	NDNPPEILQE	YVVICPKPMG	YTDILAVDPD	600
	EPVHGAFYFP	SLPNTSPEIS	RLWSLTKVND	TAARLSVQKN	AGFOEYTIPI	TVKDRAGQAA	660
	TKLLRVNLCE	CTHPQCRAT	SRSTGVILGK	WAILAILLGI	ALLPSVLLTL	VOGVFGATKG	720
	KRFPEDLAQQ	NLIISNTEAP	GDDRVCSANG	FMTQTNNSS	QGFCTMGSG	MKNGGQETIE	780
65	MMKGGNQTLE	SCRAGHHHT	LDSCRGGHTE	VDCRYTYSE	WHSFTQPRLG	EBSIRGRTG	

Seq ID NO: 528 DNA sequence
Nucleic Acid Accession #: NM_001941.2
Coding sequence: 64..2754

70	1	11	21	31	41	51	
	GGCAGGTCTC	GCTCTCGGCA	CCCTCCCGGC	GCCCGCGTTC	TCCTGGCCCT	GCCCGGCATC	60
	CGATGGCCG	CCGCTGGGCC	CCGGCGCTCC	GTGCGCGGAG	CCGTCTGCCT	GCATCTGCTG	120
	CTGACCCCTG	TGATCTTCAG	TCGTGATGGT	GAAGCCTGCA	AAAAGGTGAT	ACTTAAATGTA	180
75	CCTTCTAAAC	TAGAGGCAGA	CAAAATAATT	GGCAGAGTTA	ATTGGGAAGA	GTGCTTCAGG	240
	TCTGCAGACC	TCATCCGGTC	AAGTGATCCT	GATTTTCAGAG	TTCTAAATGA	TGGGTCAGTG	300
	TACACAGCCA	GGGCTGTGTC	GCTGTCTGAT	AAGAAAAGAT	CATTTACCAT	ATGGCTTTCT	360
	GACAAAGGA	AACAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAAGAAGGTA	420
	TGGAAGACAA	GACACACTAG	AGAACTGTGT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCACCT	480
80	ATTCCTTGCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTCC	CATTGTTTCT	TCACAAAGTT	540
	GAATCTGATG	CAGCACAGAA	CTATACTGTC	TTCTACTCAA	TAAAGTGGACG	TGGAGTTGAT	600
	AAAGAACCTT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAAATCTATT	TTGCACTCGG	660
	CCTGTGGATC	GTGAAGAATA	TGATGTTTTT	GATTTGATTG	CTTATGCGTC	AACTGCAGAT	720
	GGATATTTCG	CAGATCTGCC	CCTCCCACTA	CCCATCAGGG	TAGAGGATGA	AAATGACAAC	780
85	CACCTGTCTT	TCACAGAGC	AATTTATAAT	TTTGAAGTTT	TGGAAGTAG	TAGACCTGGT	840
	ACTACAGTGG	GGGTGGTTTG	TGCCACAGAC	AGAGATGAAC	CGACACAAT	GCATACGCGC	900
	CTGAAATACA	GCAATTTGCA	GCAGACACCA	AGGTCACTTG	GGCTCTTTTC	TGTGCATCC	960

	AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
	TCATTGATAA	TGAAGGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTGAT	AGGCACATCA	1080
	ACTTTGATCA	TAACAGTAAC	AGATTCAAAT	GATAATGCAC	CCACTTTCAG	ACAAAATGCT	1140
	TATGAAGCAT	TTGTAGAGGA	AAATGCATTG	AATGTGGAAA	TCTTACGAAT	ACCTATAGAA	1200
5	GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTIT	AAAGGGAAAT	1260
	GAAATGGAC	ATTTCAAAAT	CAGCACAGAC	AAAGAAACTA	ATGAAGGTGT	TCTTTCTGTT	1320
	GTAAGGCCAT	TGAATTATGA	AGAAAACCGT	CAAGTGAACC	TGGAATTTGG	AGTAAACAAT	1380
	GAAGGCCAT	TTGCTAGAGA	TATTCGCCAG	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
	GTTCATGTGA	GGGATCTGGA	TGAGGGGCGT	GAATGCACTC	CTGCAGCCCA	ATATGTGGCG	1500
10	ATTAAGAAAA	ACTTAGCAGT	GGGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
	AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCAAAAATG	AGTTGTATTA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAATGTATA	ATCCACCAGA	AATACTTCAA	1800
15	GAATATGTAG	TCATTGTCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCACTTTGC	CCAATACTTC	TCCAGAAATC	1920
	AGTAGACTGT	GGAGCCTCAC	CAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGAATA	TACCAITCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
	GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACCTA	GTGTCGTGCG	2100
20	ACTTCAAGTA	GTCAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCCTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTITGG	TGCAACTAAA	2220
	GGGAAACGTT	TTCTGAAGA	TTTAGCACAG	CAAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
	CCTGGAGAGC	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAACTTAC	CAACAACCTC	2340
	AGCCAAGGTT	TTTGTGGTAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAAACCAT	2400
25	GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCTGGAGCT	CCTGAGGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACTTACTCG	2520
	GAGTGGCACA	GTTTACTACA	ACCCGCTCTC	GGTGAATAAT	TGCTATGATG	TAATCAGAAT	2580
	GAAGACCCCA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGATCT	2640
	CCAGCTGGTT	CTGTGGGCTG	CTGCAGTGAA	AAGCAGGAAG	AAGATGGCCT	TGACTTTTTA	2700
30	AATAAATTTG	AAOCCAAAT	TATTACATTA	GCAGAAGCAT	GCAAAAAGAG	ATAATGTCAAC	2760
	AGTGCTACAA	TTAGGTCTTT	GTGACACATT	CTGGAGGTTT	CCAAAAATAA	TATTGTAAAG	2820
	TTCAATTTCA	ACATGTATGA	ATATGATGAT	TTTTTTCTCA	ATTTTGAATT	ATGCTACTCA	2880
	CCAATTTATA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAAATGTITAA	2940
35	AACAGACAAC	TGGTAAATCT	CAAACTCCAG	CAGTGGAAAT	AAGGTCTCTA	AAGCATCTGC	3000
	TCCTTTTITT	TTTTACGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAAAC	3060
	ATAGCTAAGT	TATGCTAATA	TCACATTATT	ATGTATTTCAC	TTTAAAGTAT	AGTTTAAAAA	3120
	ATAAACACAG	AATATTGAGT	ATCACTATGT	GAAGAAAGTT	TTGAAAAGAA	AACATGAAG	3180
	ACTGAATTAA	ATTAAAAATG	TTGCAGCTCA	TAAAGAAATG	GGACTCACCC	CTACTGCACT	3240
40	ACCAAAATCA	TTTGACTTTG	GAGGCAAAAT	GTGTGTGAAT	GCCTTATGAA	GTAGCAATTT	3300
	TCATATAGAA	TATAGTTTGA	AATAAATGTG	TGTGTGTATA	TTATTATTAA	TCAATGTCAAT	3360
	ATTTAAATGT	AAATGAGAAA	AAAGAGGAAA	ATGTGAAAAA	CTTGAAATGA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCTTAGG	CCTGGGCTCT	TAAATGTCTG	3480
	ATTATAACTG	AGTCTATGAG	GAAATAGTTC	CTGTCCAATT	TGTGTAAATT	GTTTAAAAAT	3540
45	GTAAATAAAT	TAAACTTTTC	TGGTTCTGTT	GGGAAGGAAA	TAGGGAATCC	AATGGAACAG	3600
	TAGCTTTGCT	TTGACGTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATACTCGC	TGCAGCTGGG	GTTCCTGCTC	TTTTGGTAGC	AAGGCTCCAG	AGATGAGGTG	3720
	TTTTTTTTCG	GGAGCTAATA	ACAAAAACAT	TTTAAACTTT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATTGCT	GTTCCTATTG	TCTCTTATAG	TGACCAACAT	TTTTTTAATT	TAGATCCAAA	3840
50	TAACCATGTC	CTCTAGAGTT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
	GCACCTGGGG	GAGATGTGAT	GTCTTTAAAC	CTAAGCCCCA	CAAACTTGAC	ACCTGATCAG	3960
	GTCTGGGAGC	TACAAAATTT	CATTTTCTCT	CTCACTGCCC	TTCTTCTGAG	TGGCATTGGC	4020
	CTGAATCAAG	GAAGGCCAGG	CCTTGTGGGC	CCCTTCTTTT	CGGCTTCTTG	CTAAAGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAAGTGACTC	CAGGTTTCTC	ACCATCCTTC	AGCGTGAATT	4140
55	AAITTTTAAAT	CAGTTTGCTT	TCTCCAGAGA	AAITTTTAAA	TAAATAGAAG	AATAGAAATT	4200
	TTGAATGTAT	AAAGAAAAAA	GATCAAGTTG	TCATTTTAGA	ACAGAGGGAA	CTTTGGGAGA	4260
	AAGCAGCCCA	AGTAGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCTTCAA	4320
	GGGCAAGGAG	AGGCCACAAG	GAATATGGGT	GGGAGTAAAA	GCAACATCGT	CTGCTTCATA	4380
	CTTTTTCCTA	GGCTTGGCAC	TGCCTTTTCC	TTTCTCAGGC	CAATGGCAAC	TGCCATTGTA	4440
60	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTGGGA	GTGAGAAATC	4500
	AAGGAGACAG	AGCTCACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACTG	4560
	GTGTGTCAGA	ACAAACAAGG	CATTATGGGG	AAITGTTGTA	TTCTTCTGCG	AGCCCTCCTT	4620
	CTGGGCACTA	AGAAGGTCTA	TGAATTAAT	GCCTATCTAA	AAITCTGATT	TATTCCTACA	4680
	TTTTCTGTTT	TCTAATTTGA	CCCTAAAATC	TATGTGTTTT	AGACTTAGAG	TTTTTATTGC	4740
65	CCCCCCCCCC	TTTTTTTTTG	AGACGGAGTC	TGCTCTGAC	GCACAGGCTG	GAGTGCACTG	4800
	GCTCCGATCT	CTGCTCACTG	AAAGCTCCGC	CTCCCGGTTT	CATGCCATTC	TCCTGCTCTA	4860
	GCCTCCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCAACACG	CCCGGCTAAT	TTTTTGTATT	4920
	TTTAATAGAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTCGATCTCC	TGACCTCGTG	4980
	ATCCGCTCTG	CTCGGCTCTC	CAAAAGTCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040
70	CTTGTTTTCC	GTTTAAAGTC	GTCTTCTTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAAGT	5100
	TGATCATACG	AATTTGATCA	ATCTTGAAT	ACTCAACCAA	AAGACAGTGC	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGSCA	CAAAATATTG	GTCTGAGAAAT	GGAATTCCTC	GTAAGCCTAG	5220
	TTGCTGAAAT	TTCTGCTGTT	AACCAGAAAG	CAGTTTTATC	TAACGGCTAC	TGAAACACCC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACTCACCG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
75	CTAGTGCCGA	TAAACTTTCT	CAAAAGACAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
	TAACCATCTC	TTTGTCTCTT	GAACATGCTG	AAAACCACTT	GGTCTGATG	TATGCCCGAA	5460
	TTTGTAAATC	TTTTCTCTCA	AATGAAAATT	TAAITTTAGG	GATTCAITTC	TATATTTTCA	5520
	CATATGTAGT	ATTATTATT	CCTTATATGT	GTAAGGTGAA	ATTTATGGTA	TTTGAGTGTG	5580
	CAAGAAAATA	TATTTTAAA	GCTTTTCAAT	TTCCCGCAGT	GAATGATTTA	GAATTTTTTA	5640
80	TGTAATATTA	CAGAAATGTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAAAATGCAGT	5700
	GGGGTTTGTG	TGTCAATGTT	TTAAACAGAG	TTTTAGTATT	GCTATTAATA	GAAGTTACTT	5760
	TGCTTTTAAA	GAAACTTGGC	TGCTTAAAT	AAGCAAAAAT	TUGATGCATA	AAGTAATATT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAAAAT	ATTAACCTGG	TTTCTTGT	TTGCTGTATT	5880
	TAGAGATTAA	ATAATTCTAA	GATGATCACT	TGCAAAAAT	ATGCTTATGG	CTGGCATGGA	5940
85	AATAGAAAAT	CTCAATTTATG	TCTTTTGTGT	ATTAATGGGG	AATATTTTGG	ACAAATGTTT	6000
	ATTATCAAAAT	TGTCGACATC	ATTAATATAT	ATTGTAATGT	TGGGAAGAGA	TCACTATTTT	6060
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAAT	TTGATCGGGT	6120
	ATTAAGAGTA	TTAGAGGGTG	GTTATAATTG	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180

5
10
15

```

AGGGGTTTCTA CTTTGAGGAC CAGTGTAGTC AAGGGAAAAC ATGAGTTAAA AAGAAAAGCA 6240
GGCAATATTG CAGTCTTGAT TCTGCCACIT ACAGGATAGA TAATGCTTGA ACTTTAATGA 6300
CAAGATGATC CAACCATAAA GGTGCTCTGT GCTTCACAGT GAATCTTTTC CCCATGCAGG 6360
AGTGTGCTCC CTTACAAAAG TTAAGACTGA TCATTTCAAA AATCTATTAG CTATATCAAA 6420
AGCCTTACAT TTTAATATAG GTTGAACCAA AATTTCAAIT CCAGTAACCT CTATTGTAAC 6480
CATTATTTTT GTGTATGTCT TCAAGATGT TCATTGGATT TTTGTTTGA ATAGTAAAT 6540
ACCGGATACA TTTCACGTGT CCTTCAGTAT TGATTTGGTT GAATATTGGG TCATAATGGT 6600
TGAGAAGCAT GGACACTAGA GCCAGAATGC TTGGATATGA ATCCTGGATC TGTCACTTAC 6660
TTCTGTGTGA CCTTTGAAAG GCTACTTATT TCCTCTCTTA GCTTCTCAT TAAATCAAT 6720
GAACAATGCC AGCCTCATGG GGTGTGTGAA TGATTAATAT AGTTAATATA CCTAAGTAC 6780
ATAGAACACT GCCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840
GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTT AGGGAAATAA AGTTTGTGCA 6900
TATATATAAT CCGGAACAT G

```

Seq ID NO: 529 Protein sequence
Protein Accession #: NP_001932.1

20
25
30
35

```

1 11 21 31 41 51
MAAAGPRRSV RGAVCLHLL TLVIFSRDGE ACKKVLNVN SKLEADKIIG RVNLEBCFRS 60
ADLIRSDPD FRVLNDGSVY TARAVALSDK KRSPTIWLSD KRKQTQKEVT VLLERQKKVS 120
KTRHRETLYL RRAKRRWAPI PCSMQENSLG PFPLFLQQVE SDAQNYYTVF YSISGRGVDK 180
EPLNLFYIER DTGNLFCTRP VDREEDVDFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240
FVPTAIIYNF EVLESSRPGT TVGVVCATDR DEPDIMHTRL KYSILQQTFR SPGLFSVHPS 300
TGVITTVSHY LDREVDVKYS LIMKVQDMDG QPFGLIGTST CIITVTSND NAPTRQNAV 360
EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKQNE NGHPKISTDK ETNEGVLVV 420
KPLNYEENQV NLEIGVNNIE AFFARDIPRV TALNRALVTV HVRDLDEGPE CTAAQYVRI 480
KENLAVGSKI NGYKAYDPEN RNNGNGLRYKK LBDPKGWITI DEISGSIITS KILDREVETP 540
KNELYNIYVL AIDKDRSCT GTLAVNIEDV NDNPPILQBE YVVICPKMG YTDILAVDPD 600
EPVHGAPFFY SLPNTSPEIS RLWSLTKVND TAARLSYQKN AGFQYTIPI TVKDRAGQAA 660
TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLPSVLLTL VCGVFGATGK 720
KRFPEDLAQP NLIIISNTEAP GDDRVCSSANG FMTQTNNSS QGFCGTMGSG MKNGQETIE 780
MMKGGNQTLE SCRGAGHHHT LDSCRGGHTE VNCRYTYSE WHSPQTQRLG EKLHRCNQNE 840
DRMPSQDYVL TYNVEGRGSP AGSVGCCSEK QEEDGLDPLN NLEPKFITLA RACTER

```

Seq ID NO: 530 DNA sequence
Nucleic Acid Accession #: NM_016583.2
Coding sequence: 72..842

40
45
50
55
60

```

1 11 21 31 41 51
GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60
TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCGAGA 120
CCATGGCCCA GTTTGGAGGC CTGCCCGTGC COCTGGACCA GACCCCTGCC TTGAATGTGA 180
ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAT GCCTCAGCA 240
ATGGCCCTGCT GTCTGGGGGC CTGTGGGCA TTCTGGAAA CCTTCOGCTC CTGGACATCC 300
TGAAGCCTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGA AAAGTGACGT 360
CAGTGATTCC TGGCCTGAAC AACATCATTG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420
AACTTGGCCT TGTGCAGAGC CCGTATGGCC ACCGCTCTTA TGTCAACATC CCTCTCGGCA 480
TAAAGCTCCA AGTGAATACG CCCCTGGTCG GTGCAAGTCT GTTGAGGCTG GCTGTGAAGC 540
TGGACATCAC TGCAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC 600
TTGGTGACTG CACCCATTCC CCTGGAAGCC TGCAAAATTC TCTGCTTGAT GGACTTGGCC 660
CCCTCCCATC TCAAGGCTCT CTGGACAGCC TCACAGGGAT CTGGAATAAA GTCTCTGCCTG 720
AGTTGGTTCA GGGCAACGTC TGCCCTCTGG TCAATGAGGT TCTCAGAGGC TTGGACATCA 780
CCCTGGTGCA TGACATTGTT AACATGCTGA TCCACGACT ACAGTTTGTG ATCAAGGTCT 840
AAGCCTTCCA GGAAGGGGCT GGCCTCTGCT GAGCTGCTTC CAGTGCTCA CAGATGGCTG 900
GCCCATGTGC TGGAGATGA CACAGTTGCC TTCTCTCCGA GGAACCTGCC CCCTCTCCTT 960
TCCCACGAGG CGTGTGTAAC ATCCCATGTG CCTCACCTAA TAAATGGCT CTCTCTCTGC 1020
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

```

Seq ID NO: 531 Protein sequence
Protein Accession #: NP_057667.1

65
70

```

1 11 21 31 41 51
MFQTGGLIVF YGLLAQTMAG FGGLPVPLDQ TLPLNVNPAI PLSPTGLAGS LTNALSNGLL 60
SGGLIGILEN LPLLDILKPG GGTSGGLLGG LLGKVTSVIP GLNNIIDIKV TDPQLLELGL 120
VQSPDGERLY VTIPLGILKQ VNTPLVGASL LRLAVKLDIT ABILAVRDQK ERIHLVLGDC 180
THSPGSLQIS LLDGLGPLPI QGLLDLSTGI LNKVLPELVQ GNVCLPVNEV LRGLDITLVH 240
DIVNMLIHGL QFVIKV

```

75
Seq ID NO: 532 DNA sequence
Nucleic Acid Accession #: NM_004363.1
Coding sequence: 115..2223

80
85

```

1 11 21 31 41 51
CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 60
TCCTGGAACCT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 120
TCTCCCTGGG CCGCTCCCCA CAGATGGTGC ATCCCTGGCC AGAGGCTCCT GCTCACAGCC 180
TCACTTCTAA CCTTCTGAA CCGGCCAACC ACTGCCAAGC TCACTATTGA ATCCACGCCG 240
TTCAATGTGC CAGAGGGGAA GGAGGTGCTT CTACTTGTCC ACAATCTGCC CCAGCATCTT 300
TTTGGGTACA TTTGGGTACA AGGTGAAAGA GTGATGGCA ACCGTCAAT TATAGGATAT 360
GTAATAGGAA CTCACACAGC TACCCAGGG CCGCATACA GTGGTCGAGA GATAATATAC 420
CCCAATGCAT CCCTGCTGAT CCAGAACATC ATCCAGAAATG ACACAGGATT CTACACCTTA 480

```

5
10
15
20
25
30
35
40
45

CAOCTCATAA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCCG GGTATACCCG 540
GAGCTGCCCA AGCCTCCAT CTCCAGCAAC AACTCCAAAC CCGTGGAGGA CAAGGATGCT 600
GTGGCCTTCA CCTGTGAACC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAAT 660
CAGAGCCTCC CGGTCACTCC CAGGCTGCAG CTGTCCAATG GCAACAGGAC CCTCACTCTA 720
TTCAATGTCA CAAGAAATGA CACAGCAAGC TACAAATGTG AAACCCAGAA CCCAGTGAGT 780
GCCAGGCGCA GTGATTCAGT CATCCTGAAT GTCTCTATG GCCCGGATGC CCCCAACATT 840
TCCCTCTTAA ACACATCTTA CAGATCAGGG GAAATCTGA ACCTCTCCTG CCACGCGAGC 900
TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTCTCCAGCA ATCCACCCAA 960
GAGCTCTTTA TCCCAACAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT 1020
AACTCAGACA CTGGCCTCAA TAGGACCACA GTCACGACGA TCACAGTCTA TGCAGAGCCA 1080
CCCAAAACCT TCATCACCAG CAACAACCTC AACCCCGTGG AGGATGAGGA TGCTGTAGCC 1140
TTAACTGTG AACCTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC 1200
CTCCCGGTCA GTCCAGGCT GCAGCTGTCC AATGACAA CA GACCCCTCAC TCTACTCAGT 1260
GTCACAAGGA ATGATGTAGG ACCCTATGAG TGTGGAATCC AGAACGAATT AAGTGTGAC 1320
CACAGCGACC CAGTCATCTT GAATGCTCTT TATGGCCCGC AGGACCCAC CATTTCCTCC 1380
TCATACACCT ATTACCGTCC AGGGGTGAAC CTCAGCCTCT OCTGCCATGC AGCCTCTAAC 1440
CCACCTGCAC AGTATCTTGG GCTGATTGAT GGGAAACATCC AGCAACACAC ACAAGAGCTC 1500
TTTATCTCCA ACATCACTGA GAAGAACAGC GGACTCTATA CCTGCCAGGC CAATAACTCA 1560
GCCAGTGGCC ACAGCAGGAC TACAGTCAAG ACAATCACAG TCTCTGCGGA GCTGCCCAAG 1620
CCCTCCATCT CTGCAACAA CTCCAAACCC GTGGAGGACA AGGATGCTGT GGCCTTCACC 1680
TGTGAACCTG AGGCTCAGAA CACAACCTAC CTGTGGTGGG TAAATGGTCA GAGCCTCCCA 1740
GTCACTGCA GCGCTCAGCT GTCCATGGC AACAGGACCC TCACTCTATT CAATGTCA 1800
AGAAATGACG CAAGAGCCTA TGTATGTGGA ATCCAGAACT CAGTGAATGC AAACCGCAGT 1860
GACCCAGTCA CCTGTGATGT CCTCTATGGG CCGGACACCC CCATCATTTT CCCCCAGAC 1920
TCGTCTTACC TTTCCGGAGC GAACCTCAAC CTCTCCTGCC ACTCGGCTC TAACCCATCC 1980
CCGCACTATT CTTCGGGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTATC 2040
GCCAAATCA CGCCAAATTA TAACGGGACC TATGCTCTGT TTGTCTCTAA CTGGCTACT 2100
GGCCGCAATA ATTCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TCTCTCTGT 2160
CTCTCAGCTG GGGCCATGT CGGCATCATG ATTGGAGTGC TGGTGGGGT TGCTCTGATA 2220
TAGCAGCCTG GGTGTAGTTT CTTCATTCA GGAAGACTGA CAGTTGTTTT GCTTCTCTCT 2280
TAAAGCATTT GCAACAGCTA CAGTCTAAAA TTGCTCTTT ACCAAGGATA TTTACAGAAA 2340
AGACTCTGAC CAGAGATCGA GACCATCTTA GCCAACATCG TGAAACCCCA TCTCTACTAA 2400
AAATACAAAA ATGAGCTGGG CTGTGTGGCG CGCACCTGTA GTCCAGTTA CTCGGGAGGC 2460
TGAGGCGAGA GAATGCTGTG AACCCGGGAG GTGGAGATTG CAGTGAGCCC AGATCGCACC 2520
ACTGCACCTC AGTCTGGCAA CAGAGCAAGA CTCCATCTCA AAAAGAAAAG AAAAGAAGAC 2580
TCTGACCTGT ACTCTGAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTTCCAA 2640
AACTTAATG AACTAACTGA CAGCTTCATG AAACGTGCCA CCAAGATCAA GCAGAGAAAA 2700
TAATTAATTT CATGGGACTA AATGAACATA TGAGGATTGC TGATTCTTTA AATGTCTTGT 2760
TTCCCGATT TTAGGAAACT TTTTCTCTT TAAGCTATCC ACTCTTACAG CAATTTGATA 2820
AAATATACTT TTGTGAACAA AAATTGAGAC ATTTACATTT TCTCCTATG TGGTGGCTCC 2880
AGACTTGGGA AACTATTCAT GAATATTAT ATTGTATGGT AATATAGTTA TTGCACAAGT 2940
TCAATAAAAA TCTGCTCTTT GTATAACAGA AAAA

Seq ID NO: 533 Protein sequence
Protein Accession #: NP_004354.1

50
55
60

1 11 21 31 41 51
MESPSAPPHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVABGKE VILLVHNLPO 60
HLFGYSWYRG ERVDGNRQII GYVIGTQOAT PGPAYSGREI IYPNASLLIQ NIIQNDTGFP 120
TLHVIKSDLV NEEATGQFRV YPELPKPSIS SNNSKPVEDK DAVAPTCEPE TQDATYLMWV 180
NNQSLVPSPR LQLSNGNRTL TLFNVTRNDT ASYKCEIQNP VSARRSDSVI LNVLYGPDAP 240
TISPLNTSYR SGENLNLSCB AASNPQAQYS WFWNGTFQQS TQELFIPNIT VNNSGSYTCQ 300
AHNSDTGLNR TTVTTITVYA EPPKPPITEN NSNPVEDEDA VALTCEPEIQ NTTYLMWVNN 360
QSLFVSPRIQ LSNDRNLTL LSVTRNDVGP YECGIQNELS VDHSDPVLN VLYGDDPTI 420
SPSYTYRRG VNLSLSCHAA SNPPAQYSWL IDGNIQHQRT ELFIENITEK NSGLYTCQAN 480
NSASGHSRRT VKTITVSAL PKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLWVWNGQS 540
LPVSPRLQLS NGKRLTLFNF VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDPTIISP 600
FDSSYLSGAN LNLSCHSASN PSPQYSWRIN GIPQHQTVL FIAKITPNMN GTYACFVSNL 660
ATGRNNSIVK SITVSAGTS PGLSAGATVG IMIGVLVGVA LI

Seq ID NO: 534 DNA sequence
Nucleic Acid Accession #: NM_006952.1
Coding sequence: 11..793

65
70
75
80
85

1 11 21 31 41 51
AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTTCGTTGC TTCCAGGGCC TGCTGATTTT 60
TGGAATGTG ATATTGGTT GTTGCAGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTGT 120
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACACGATG ACATCTATGG 180
GGCTGCCTCG ATCGGCATAT TTGTGGGCAT CTGCCTCTTC TGCCGTGCTG TTTCTAGGCAT 240
TGTAGGCATC ATGAAGTCCA GCAGGAAAA TCTTCTGGCG TATTTCTATC TGATGTTTAT 300
AGTATATGCC TTTGAAGTGG CATCTGTGAT CACAGCAGCA ACACAAAGAG ACTTTTTCAC 360
ACCCAACCTC TTCTGAAGC AGATGCTAGA GAGGTACCAA AACACAGGCC CTCCAAACAA 420
TGATGACCA TGGAAAAACA ATGGAGTCAC CAAACCTGG GACAGGCTCA TGCTCCAGGA 480
CAATTGCTGT GCGCTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CCTTCCGGAC 540
TGAGAATAAT GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA 600
AGAACCTCTC AACCTGGAGG CTGTGTAACCT AGGCGTGCCT GGTTTTTATC ACAATCAGGG 660
CTGCTATGAA CTGATCTCTG GTCCAAATGAA CCGACACGCC TGGGGGGTTG CTGTGTTGG 720
ATTGCTCATT CTCTGCTGGA CTTTGTGGGT TCTCCTGGGT ACCATGTTCT ACTGAGCAG 780
AATTGAATAT TAAGAA

Seq ID NO: 535 Protein sequence
Protein Accession #: NP_008883.1

85

1 11 21 31 41 51

MAKDNSTVRC FQGLLIPGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAW 60
 IGIFVGLICLF CLSVLGLVGI MKSSRKILLA YFILMFIYVA FEVASCITAA TQRDFPTFNL 120
 FLKQMLERYQ NNSFPNNDQ WKNNGVTKTW DRLMLQDNCC GVNGPSDWQK YTSAFRTENN 180
 DADYPWPRQC CVMNKLKEPL NLEACKLGVP GFYHNQGCYE LISGPMNRHA WGVANFGFAI 240
 LCWTFWVLLG TMFYWSRIEY

Seq ID NO: 536 DNA sequence
 Nucleic Acid Accession #: NM_002638.1
 Coding sequence: 120..473

1 11 21 31 41 51
 CAATACAGCT AAGGAATTAT CCCTTGTAAG TACCACAGAC CGGCCCTGGA GCCAGGCCAA 60
 GCTGGAGCC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
 TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGCTGGGACG CTGGTTCTAG 180
 AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCCTA 240
 TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAATTAA AGGTCAAGAT AAAGTCAAAG 300
 CGCAAGAGCC AGTCAAAGGT CCAAGTCTCCA CTAAGCCTGG CTCTGCCCC ATTATCTTGA 360
 TCCGGTCCGC CATGTGGAAT CCCCTAACC GCTGCTTGAA AGATACTGAC TGCCCAAGAA 420
 TCAAGAAGTG CTGTGAAGGC TCTTGCGGGA TGGCCTGTTT CGTTCCCCAG TGAAGGGAGC 480
 CGGTCTCTGC TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCATCTGGT CCTAAGTCCC 540
 TGCTGCCCTT CCCCTTCCA CACTGTCCAT TCTTCTCCC ATTACAGATG CCCACGGCTG 600
 GAGCTGCCCT TCTCATCCAC TTTCCAATAA A

Seq ID NO: 537 Protein sequence
 Protein Accession #: NP_002629.1

1 11 21 31 41 51
 MRASSFLIVV VFLIAGTLVL EAAVTGVFVK GQDTVKGRVP FNGQDPVKQ VSVKQGDVKV 60
 AQEFVKGPVS TKGPSCPIIL IRCAMLNPPN RCLKDTDCPG IKKCEBSCG MACFVPG

Seq ID NO: 538 DNA sequence
 Nucleic Acid Accession #: NM_001793.2
 Coding sequence: 71..2560

1 11 21 31 41 51
 AAAGGGGCAA GAGCTGAGCG GAACACCGGC CGGCCCTGCG GGCAGCTGCT TCACCCCTCT 60
 CTCTGCAGCG ATAGGGCTCC CTGCTGGACC TCTCGCTGCT CTCTCCTTC TCCAGGTTTG 120
 CTGGCTGACG TGCGCGGCGT CCGAGCCGTG CCGGGCGGTC TTCAGGGAGG CTGAAGTGAC 180
 CTTGGAGGCG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAAG TATTATGCGG 240
 CTGCCCTGGG CAAGAGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300
 TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360
 ATCCAAACGT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420
 TGAAGATGCG AAGGGTCCCT TCCCCAGAG ACTGAATCAG CTCAGTCTA ATAAAGATAG 480
 AGACACCAAG ATTTTCTACA GCATCACGGG CGCGGGGSCA GACAGCCCC CTGAGGGTGT 540
 CTTGCTGTGA GAGAAGGAGA CAGGCTGGTT GTTGTGTAAT AAGCCACTGG ACCGGGAGGA 600
 GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTACAGG AATGTTGCT CAGTGGAGGA 660
 CCCCATGAAC ATCTCCATCA TGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720
 GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCTACCA GGTACTTCTG TGTATCAGGT 780
 GACAGCCAGG GATGAGGATG ATGCCATCTA CACTACAAT GGGGTGGTTG CTTACTCCAT 840
 CCTATGCCAA GAACCAAGG ACCCACACGA CCTCATGTTT ACCATTACCC GGAGCACAGG 900
 CACCATCAGC GTCATCTCCA GTGGCCTGGA CCGGGAAGAA GTCCCTGAGT ACACACTGAC 960
 CATCCAGGCC ACAGACATGG ATGGGGACGG CTCCACCAAC AOGGCAGTGG CAGTAGTGGA 1020
 GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080
 GCCTGAGATG GCAGTGGGCT ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140
 CAACTCACCA GCGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACG GGGACCATTT 1200
 TACCATCACC ACCCACCTTG AGAGCAACCA GGGCATCTGT ACAACAGGGA AGGGTTTGGA 1260
 TTTTGAAGCC AAAAACCCAG ACACCCCTGT CGTTGAAGTG AOCACAGGAG CCCCTTTTGT 1320
 GCTGAAGCTG CCAACCTCCA CAGCCACCAT AGTGTGCCAG GTGGAGGATG TGAATGAGGC 1380
 ACCCTGTGTT GTCCACCTT CCAAAGTCGT TGAGTCCAG GAGGGCATCC CCACTGGGGA 1440
 GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCAGCTACCG 1500
 CATCTTGAGA GACCCAGCAG GGTGGCTAGC CATGGAACCA GACAGTGGGC AGGTACAGC 1560
 TGTGGGCAAC CTGACCCGTG AGGATGAGCA GTTTGTGAGG AACACATCT ATGAAGTCAT 1620
 GGTCTTGGCC ATGGACAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGTAAAC 1680
 ACTGATTGAT GTCAATGACC ATGGCCAGT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740
 CCAAAGCCCT GTGGCCAGAG TGCTGAACAT CACGACAAAG GACCTGTCTC CCCACACTC 1800
 CCCCTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGAACGGCAG AGGTCAACGA 1860
 GGAAGGTGAC ACATGGTGTCT TGTCCCTGAA GAAGTTCCTG AAGCAGGATA CATATGACGT 1920
 GCACCTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980
 GTGCGACTGC CATGGCCATG TCGAAACCTG CCCTGACCCC TGAAGGGAG GTTTCATCCT 2040
 CCTGTGCTGC GGGGCTGTCC TGCTCTGTCT GTTCTCTCTG CTGGTGTCTG TTTTGTGGT 2100
 GAGAAAGAG CGGAAGATCA AGGAGCCCTT CCTACTCCCA GAAGATGACA CCGTGTACAA 2160
 CGTCTTCTAC TATGGCGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCAACCA 2220
 GCTCCACCGA GGTCTGGAGG CAGGCGCGGA GGTGTTCTCT CGCATGAAG TGGCAACAA 2280
 CATCATCCCG ACACCATATG ACCGTCTCTG GCCAGCCAAC CCAGATGAAA TCGGCAACTT 2340
 TATAATTGAG AACTTGAAGG CGGCTAACAC AGACCCCAAC GCCCGGCCCT ACAGACACCT 2400
 CTTGGTGTTC GACTATGAGG GCAGCGGCTC CGACCGCGCG TCCCTGAGCT CCCTCACCTC 2460
 CTCGCGCTCC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCGCTTCAA 2520
 GAAGCTGGCA GACATGTACG GTGGCGGGGA GGACGACTAG GCGGCTGCCC TGCAGGGCTG 2580
 GGGACCAAC GTGAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG 2640
 GACTTGGGAG CTTGTGAGGA AGTGGCGGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700
 ACGTTAGAGT GGTGTCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTGTGACTC 2760
 AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCTCTAG AGGCCAAGTT TCCAGAGCC 2820
 TCTTACCTGC CGTAAATATG TCAACCTGT GTCTGGGCC TGGGCTGCT GTGACTGACC 2880
 TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTTT 2940

TTTTTTTAAAT GCTATCTTCA AAACGTTAGA GAAAGTCTT CAAAAGTGCA GCCCAGAGCT 3000
GCTGGGCCCA CTGGCCGTCC TGCATTCTG GTTTCAGAC CCCAATGCCT CCCATTCGGA 3060
TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGGTTGCC CCTTATTTT TATTTTCCCT 3120
GTTGGGTGCG TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTTTAT 3180
TAAAGAAACT TTTCACAGAA AAAAA

Seq ID NO: 539 Protein sequence
Protein Accession #: NP_001784.2

1 11 21 31 41 51
MGLFRGPLAS LLLQLQCNLQ CAASEPCRAV FREAEVTLEA GGABQEPGQA LGKVPNGCPG 60
QEPALFSTDN DDFTVRNGET VOERRSLKER NPLKIPPSKR ILRRHKRDWV VAPISVPENG 120
RGFPFQRLNQ LKSNKDRDRTK IFYSITGPGA DSPPEGVFAV EKETGWLILLN RPLDREBIK 180
YELFGHVAVE NGASVEDPMN ISIIVTDQND HKPKFTQDTP RGSVLEGVLP GTSVMQVTAT 240
DEDDATYTYN GVVAYSIHQS RPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300
TMDGDSGTTT TAVAVVEILD ANDNAPMDFP QKYEAHVPEN AVGHEVQLRT VTDLDAPNSP 360
AMRATLYLNG GDDGDHFTIT THPESNQIIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420
PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGERPVC VYTAEDPDKE NQKISYRILR 480
DPAGWLAMPD DSGQVTAAGT LDREDEQFVR NNIEVVMVLA MDNGSPPTTG TGTLLLTLLD 540
VNDHGVVPEP RQITICNQSP VRQVLNITDK DLSPTSPFPQ AQLTDDSDIY WTAEVNNEGD 600
TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVDCD HGHVETCPGP WKGGFILPVL 660
GAVLALLPLL LVLLLVVRKK RKIKPEPLLL EDDTRDNVVF YGEGGGGEED QDYDITQLHR 720
GLEARPEVVL RNDVAPTIIP TPMVPRPAN PDEIGNFIE NLKAANTDPT APPYDTLLVP 780
DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

Seq ID NO: 540 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..672

1 11 21 31 41 51
ATGAGGCTCC AAAGACCCCG ACAGGCCCGG CGGGGTGGGA GCGCGCGGCC CCGGGGCGGG 60
CGGGGCTCCC CCTACCGGCC AGACCCGGGG AGAGGCGCGC GGAGGCTGCG AAGGTTCCAG 120
AAGGCGGGGG AGGGGGCGCC GCGCGCTGAC CCTCCCTGGG CACCGCTGGG GACGATGGCG 180
CTGCTCGCCT TGCTGCTGCT CGTGGCCCTA CCGCGGGTGT GGACAGACGC CAACCTGACT 240
GCGAGACAAC GAGATCCAGA GGACTCCAG CGAACGGAOG AGGGTGACAA TAGAGTGTGG 300
TGTCATGTTT GTGAGAGAGA AAACACTTTC GAGTGCCAGA ACCCAAGGAG GTGCAAAATGG 360
ACAGAGCCAT ACTGCGTTAT AGCGCGCGTG AAAATATTTT CACGTTTTTT CATGTTTGGC 420
AAGCAGTCTC CCGCTGGTTG TGCAGCGATG GAGAGACCCA AGCCAGAGGA GAAGCGGTTT 480
CTCTCGGAAG AGCCCATGCC CTCTTTTAC CTCAGTGTGT GTAAAATTCG CTACTGCAAT 540
TTAGAGGGGC CACCTATCAA CTCATCAGTG TTCAAAGAAT ATGCTGGGAG CATGGGTGAG 600
AGCTGTGGTG GCGTGTGGCT GGCCATCTTC CTGCTGCTGG CCTCCATTGC AGCCGGCCTC 660
AGCCTGCTTT GA

Seq ID NO: 541 Protein sequence
Protein Accession #: Eos sequence

1 11 21 31 41 51
MRLQRPRQAP AGGRRAPRG RGSPYRDPFG RGARRLRRFQ KGEGGAPRAD PFWAPLGTMA 60
LLALLLVVAL PRVITDANLT ARQRPEDSDS RTDEGDNRVW CHVCERENTP EQCNERRCKW 120
TEPYCVIAAV KIFPRFFMVA KQCSAGCAAM ERPKPEEKRF LLEEMPFFFY LKCKKIRYCN 180
LEGPPINSSV FKEYAGSMGE SCGGLWLAIL LLLASIAAGL SLS

Seq ID NO: 542 DNA sequence
Nucleic Acid Accession #: XM_035292.2
Coding sequence: 53..1576

1 11 21 31 41 51
GCTCGCTGGG CCGCGGCTCC CGGGTGTCCC AGGCCCGGCC GGTGCGCAGA GCATGGCGGG 60
TGCGGGCCCG AAGCGGCGCG CGCTAGCGGC GCCCGCGGCC GAGGAGAAGG AAGAGGCGCG 120
GGAGAAGATG CTGGCCGCCA AGAGCGCGGA CGGCTCGGCG CCGGCAGGCG AGGGCGAGGG 180
CGTGACCCCTG CAGCGGAACA TCACGCTGCT CAACGCGGTG GCCATCATCG TGGGGACCAT 240
TATCGGCTCG GGCATCTTCG TGACGCCCAC GGGCGTGCTC AAGGAGGCGG GCTCGCGGGG 300
GCTGGGCTCG GTGGTGTGGG CCGCGTGGCG CGTCTTCTCC ATCGTGGGCG CGCTCTGCTA 360
CGCGGAGCTC GGCACCAACA TCTCCAAATC GGGCGGCGAC TACGCTTACA TGCTGGAGGT 420
CTACGCTCG CTGCGCGCCT TCCTCAAGCT CTGGATCGAG CTGCTCATCA TCGCGCCTTC 480
ATCGCAGTAC ATCGTGGGCC TGCTCTTCGC CACCTACCTG CTCAGCGCGC TCTTCCCCAC 540
CTGCGCGGTG CCGGAGGAGG CAGCCAAGCT CGTGGCGTGC CTCTGCGTGC TGCTGCTCAC 600
GGCCGTGAAC TGCTACAGCG TGAAGGCGGC CACCCGCGTC CAGGATGCCT TTGCGCGCGC 660
CAAGCTCCCTG GCCCTGGCCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GAAAGGGTGA 720
TGCTGCAAT CTAGATCCCA ACTTCTCATT TGAAGGCACC AAACCTGGATG TGGGGAACAT 780
TGCTGTGGCA TTATACAGCG GCCTCTTTCG CTATGGAGGA TGGAAATTA TGAATTTCTG 840
CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCTCG GCCATCATCA TCTCCTGCC 900
CATCGTGACG CTGGTGTACG TGCTGACCAA CCTGGCTTAC TTACCAACCC TTCCACCGA 960
GCAGATGCTG TCGTCCGAGG CGTGGCGCGT GGACTTCGGG AACTATCACC TGGCGTCAT 1020
GTCTCGGATC ATCCCGGTCT TCGTGGGCGT GTCTCTGCTC GGCTCCGTCA ATGGGTCCCT 1080
GTTACATACC TCCAGGCTCT TCTTCTGGGG GTCCCGGAAA GSCCACTGCG CCTCCATCCT 1140
CTCCATGATC CACCAACAGC TCCTCACCCC CGTGGCGTCC CTGCTGTTCG CGTGTGTGAT 1200
GACGCTGCTC TACGCTTCTT CCAAGGACAT CTCTCTCGTC ATCACTTCTT TCAGCTTCTT 1260
CAACTGGCTC TGGTGGGCCG TGGCCATCAT CGGCATGATC TGGCTGCGCC ACAGAAAGCC 1320
TGAGCTTGAG CGGCCATCA AGGTGAACCT GGCCCTCGCT GTGTCTTCTA TCTTGGCCTG 1380
CCTCTTCTCG TTGCGGCTCT CTTCTGGAAA GACACCGGTC GAGTGTGGCA TCGGCTTAC 1440
CATCATCTCT AGCGGGCTGC CGTCTACTT CTTCGGGCTC TGGTGGAAAA ACAAGCCCAA 1500
GTGGCTCCTC CAGGCGCATCT TCTCCACGAC CGTCTCTGTG CAGAAGCTCA TCGAGGTGCT 1560

CCCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

Seq ID NO: 543 Protein sequence
Protein Accession #: XP_035292.2

5
10
15

1	11	21	31	41	51	
MAGAGPKRRA	LAAPAAEKEE	EAREKMLAAK	SADGSAPAGE	GEGVTIQRNI	TLINGVAIIIV	60
GTIIGSGIFV	TPTGVLEKAG	SPGLALVVWA	ACGVFSIVGA	LCYAEIGTTI	SKSGGDYAYM	120
LEVYGSIPAP	LKLMIELLII	RPSSQYIVAL	VFATYLLKPL	PPTCPVPEEA	AKLVACLCLVL	180
LLTAVNCYSV	KAATRVQDAF	AAAKLLALAL	IILIGFVQIG	KGDVSNLDPN	FSFEGTKLDV	240
GNIVLALYSG	LPAYGGWNYL	NFVTEEMINP	YRNLPALIII	SLPIVTLVYV	LTNLAYFTTL	300
STEQMLSSSA	VAVDFNGNYL	GVMSWIIPVF	VGLSCFGSVN	GSLFTSSRLP	FVGSREGHLP	360
SILSMIHPQL	LTPVPSLVFT	CVMTLLYAFS	KDIFSVINFP	SFFNWLCLAL	AIIGMIWLRH	420
RKPELERPIK	VNLALPVFFI	LACLFLIAVS	FWKTPVECGI	GPTIILSGLP	VYFFGVWVKN	480
KPKNLLQGIF	STTVLCQKLM	QVVPQET				

Seq ID NO: 544 DNA sequence
Nucleic Acid Accession #: NM_005268.1
Coding sequence: 168..989

20
25
30
35
40
45

1	11	21	31	41	51	
TAAAAAGCAA	AAGAATTCGC	GGCCGCGTCG	ACACGGGCTT	CCCGAAAAC	CTTCCCGCT	60
TCTGGATATG	AAATTCAGC	TGCTTGCTGA	GTCTTATTCG	CGGCTGCTGG	GAGCCAGGAG	120
AGCCCTGAGG	AGTAGTCAC	CAGTAGCAGC	TGACGCGTGG	GTCCACCATG	AACTGGAGTA	180
TCCTTGAGGG	ACTCTGAGT	GGGGTCAACA	AGTACTCCAC	AGCCTTTGGG	CGCATCTGGC	240
TGTCTCTGCT	CTTCATCTTC	CGGTGCTGG	TGTACTCTGT	GACGGCCGAG	CGTGTGTGGA	300
GTGATGACCA	CAAGGACTTC	GACTGCAATA	CTCGCCAGCC	CGGCTGCTCC	AACCTCTGCT	360
TTGATGAGTT	CTTCCCTGTG	TCCCATGTGC	GCCTCTGGGC	CCTGCAGCTT	ATCCTGGTGA	420
CATGCCCTCT	ACTGCTCGTG	GTCAATGACG	TGGCCTACCG	GGAGGTTTCA	GAGAAGAGGC	480
ACCGAGAGGC	CCATGGGGAG	AACAGTGGGC	GCCTCTACCT	GAACCCCGGC	AAGAAGCGGG	540
GTGGGCTCTG	GTGACATAT	GTCTGCAGCC	TAGTGTTCAC	GGCGAGCGTG	GACATCGCCT	600
TTCTCTATGT	GTTCACCTCA	TTCTACCCCA	AATATATCCT	CCCTCCTGTG	GTCAAGTGCC	660
ACGCAGATCC	ATGTCCCAAT	ATAGTGGACT	GCTTCATCTC	CAAGCCCTCA	GAGAAGAAC	720
TTTTCAACCT	CTTCATGGTG	GCCACAGCTG	CCATCTGCAT	CCTGCTCAAC	CTCGTGGAGC	780
TCATCTACCT	GGTGAGCAAG	AGATGCCAOG	AGTGCCTGGC	AGCAAGGAAA	GCTCAAGCCA	840
TGTGCACAG	TCATCACCCC	CACGGTACCA	CCTCTTCCCT	CAACAAGAC	GACCTCCTTT	900
CGGGTGACCT	CATCTTCTGT	GGCTCAGACA	GTCACTCTCC	TCTCTTACCA	GACCGCCCCC	960
GAGACCATGT	GAAGAAAACC	ATCTTGTGAG	GGGCTGCCTG	GACTGTGCTG	GCAGGTTGGG	1020
CCTGGATGGG	GAGGCTCTAG	CATCTCTCAT	AGGTGCAACC	TGAGAGTGGG	GGAGCTAAGC	1080
CATGAGGTAG	GGGCAGGCAA	GAGAGAGGAT	TCAGACGCTC	TGGAGGCCAG	TTCTAGTCC	1140
TCAACTCCAG	CCACCTGCCC	CAGCTCGACG	GCACTGGGCC	AGTTCCCCCT	CTGCTCTGCA	1200
GCTCGGTTTC	CTTTTCTAGA	ATGGAAATAG	TGAGGGCCAA	TGC		

Seq ID NO: 545 Protein sequence
Protein Accession #: NP_005259.1

50
55

1	11	21	31	41	51	
MNWSIFEGLL	SGVNKYSTAF	GRWLSLVFI	FRVLVYLVT	ERVWSDHDKD	FDCNTRQPGC	60
SNVCFDEFPF	VSHVRLWALQ	LILVTCPSLL	VVMHVAYREV	QEKHREAHG	ENSGRLVLYNP	120
GKKRGLWMT	YVCSLVFKAS	VDIAFLYVFH	SFPYKYLILP	VVKCHADPCP	NIVDCFISKP	180
SEKNIFTLM	VATAAICILL	NLVELIYLV	KRCHECLAAR	KAQAMCTGHH	PHGTTSSCKQ	240
DDLSSGDLIF	LGSDSHPPLL	PDRPRDHVKK	TIL			

Seq ID NO: 546 DNA sequence
Nucleic Acid Accession #: NM_002391.1
Coding sequence: 26..457

60
65
70
75

1	11	21	31	41	51	
CGGGCGAAGC	AGCGCGGGCA	GCGAGATGCA	GCACCGAGGC	TTCTCTCTCC	TCACCTCTCT	60
CGCCCTGCTG	GCGCTCACCT	COGCGGTGCG	CAAAAAGAAA	GATAAGGTGA	AGAAGGGCGG	120
CCCGGGGAGC	GAGTGCGCTG	AGTGGGCCTG	GGGGCCCTGC	ACCCCCAGCA	GCAAGGATTG	180
CGGCGTGGGT	TTCCGCGAGG	GCACTGCGGG	GGCCAGACCC	CAGCGCATCC	GGTGAGGGGT	240
GCCCTGCAAC	TGGAAGAAGG	AGTTTGAGGC	CGACTGCAAG	TACAAGTTTG	AGAAGTGGGG	300
TGCGTGTGAT	GGGGGCACAG	GCAACCAAGT	CCGCAAGGCC	ACCCCTGAAG	AGGCGCGCTA	360
CAATGCTCAG	TGCCAGGAGA	CCATCCGCGT	CACCAAGCCC	TGCAACCCCA	AGACCAAGGC	420
AAAGGCCCAA	GCCCAAGAA	GGAAGGGAAA	GGACTAGACG	CCAAGCCTGG	ATGCCAAGGA	480
GCCCCGTGGT	TCACATGGGG	CCTGGCCACG	CCCTCCCTCT	CCCAGGCCCG	AGATGTGACC	540
CACCAAGTCC	TTCTGTCTGC	TGTTAGCTT	TAATCAATCA	TGCCCTGCCT	TGTCCTCTCT	600
ACTCCCCAGC	CCACCCCTTA	AGTGCCCAAA	GTGGGGAGGG	ACAAGGGATT	CTGGGAAGCT	660
TGAGCCTCCC	CCAAAGCAAT	GTGAGTCCCA	GAGCCCGCTT	TTGTTCTTCC	CCACAATTCC	720
ATTACTAAGA	AACACATCAA	ATAAAGTGAC	TTTTTCCCCC	CAATAAAGC	TCTTCTTTTT	780
TAATAT						

Seq ID NO: 547 Protein sequence
Protein Accession #: NP_002382.1

80
85

1	11	21	31	41	51	
MOHRGFLLLT	LLALLALTS	VAKKDKVKK	GGPGSECAEW	AWGPCTPSSK	DCGVGFREGT	60
CGAQTQRIKC	RVPCNWKKEP	GADCKYKFEN	WGACDGGTGT	KVRQGTLLKA	RYNAQCQETI	120
RVTKPCTPKT	KAKAKARKGK	GKD				

Seq ID NO: 548 DNA sequence

Nucleic Acid Accession #: NM_006783.1
Coding sequence: 1..786

```
5 1 11 21 31 41 51
| | | | | |
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAACACTC CACCAGCATC 60
GGGAAGGTGT GGATCACAGT CATCTTTATT TTCGAGTCA TGATCCTAGT GGTGGCTGCC 120
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACOGGGATGC 180
AAAAATGTGT GCTATGACCA CTTTTCCCG GTGTCCCACTA TCCGCTGTG GGCCTCCAG 240
10 CTGATCTTCG TCTCCACCCC AGCGCTGCTG GTGCCATGC ATGTGGCCTA CTACAGGCAC 300
GAAACCACTC GCAAGTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360
ATTAATAAGC ACAAGGTTCG GATAGAGGGG TCGCTGTGGT GGACGTACAC CAGCAGCATC 420
TTTTCCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACAATGGG 480
TACCACCTGC CCTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCAACCT TGTGACTGC 540
15 TTTATTCTTA GGCCACAGA GAAGACCGTG TTTACCATTT TTATGATTC TGCCTCTGTG 600
ATTTCATGCT TCGTTAAAGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTTAGG 660
AGATCAAGA GAGCAGACAG CCAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720
CAGATGAAA TGAATGAGCT GATTTAGAT AGTGGTCAA ATGCAATCAC AGGTTTCCCA 780
20 AGCTAA
```

Seq ID NO: 549 Protein sequence
Protein Accession #: NP_006774.1

```
25 1 11 21 31 41 51
| | | | | |
MDWGLHTFI GGVNKHSTSI GKVVITVIFI FRVMILVVAA QEVWGEQED FVCNTLQPGC 60
KNVCYDHPFF VSHIRLWALQ LIFVSTPALL VAMHVAYYRH ETTRKPRRGE KRNDFKDIED 120
IKKHKVRIGS SLWNTYSSSI PFRIIFEAAP MYVPYPLYNG YHLPWVLKCG IDPCPNLVDC 180
30 FISRPTEKTV FTIFMISASV ICMLLNVAEL CYLLLLKVCPR RSKRAQTQKN HPNHALKESK 240
QNEWNEILISD SGQNAITGFP S
```

Seq ID NO: 550 DNA sequence
Nucleic Acid Accession #: NM_002571.1
Coding sequence: 99..587

```
35 1 11 21 31 41 51
| | | | | |
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCTCCTGC 60
TCACCTCTGG GGTGCCCTGT GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120
AGGACCTGGA GCTCCCAAGG TTGGCAGGGA CTTGGCACTC CATGGCCATG GGGACCAACA 180
ACATCTCCCT CATGGCGACA CTGAAGGCCCT CTCTGAGGGT CCACATCACC TCATCTGTGC 240
CCACCCCGGA GGCAACCTG GAGATCGTTC TGACACAGATG GGAGAACAAAC AGCTGTGTGT 300
AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG 360
TGGCGAAGCA GGCCACGCTG CTCGATACTG ACTACGACAA TTTCTGTGTT CTCTGCCTAC 420
45 AGGACACCCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CTTGGCCAGA GTCTCTGGTG 480
AGGACGATGA GATCATGCA GGAATTCATCA GGGCTTTTCA GCGCTGCCCC AGGCACCTAT 540
GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCGGTGCCG TTTCTAGCTC ACCTCGGCT 600
CCAGGAAGAC CAGACTCCCA CCTTCCACA CCTCCAGAGC AGTGGGACTT CCTCTGCCCC 660
50 TTTCAAGAA TAACCAAGC TCAGAAGACG ATGACGTGGT CATCTGTGTC GCCATCCCCT 720
TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCTCGGGGGC AGAGTCTCTG 780
GCAGAGGTTA TTAATAAACC CTGGAGCAT G
```

Seq ID NO: 551 Protein sequence
Protein Accession #: NP_002562.1

```
55 1 11 21 31 41 51
| | | | | |
MDIPQTKQDL ELPLAGTWH SMAMATNNIS LMATLKAPLR VHITSLLPTP EDNLEIVLHR 60
WENNSVEVK VLGEKTNPK KFKINYTVAN EATLLDLYD NFLFLCLQDT TPIQSMCCQ 120
YLARVLVEDD SIMQGFIRAF RPLPRHLWYL LDLRQMEEPC RF
```

Seq ID NO: 552 DNA sequence
Nucleic Acid Accession #: NM_006500.1
Coding sequence: 27..1967

```
65 1 11 21 31 41 51
| | | | | |
ACTTGGGTCT GCGCCTCCGG CCAAGCATGG GCGTCCCGAC GCTGGTCTGC GCCTTCTTGC 60
TGGCGGCTGT CTGTGCTGT CTCTGGGTGG GCGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
CGCTGAGCT GGTGGAGGTG GAAAGTGGCA GCACAGCCCT TCTGAAGTGC GGCCTCTCCC 180
AGTCCCAAGG CAACCTCAGC CATGTGCACT GGTTTTCTGT CCACAAGGAG AAGCGGACGC 240
TCATCTTCGG TGTGGCCAG GGCAGGGGCC AGAGCGAACC TGGGAGTAC GAGCAGCGGC 300
TCAGCCTCCA GGACAGAGGG GCTACTCTGG CCTGACTCA AGTCACCCCC CAGACGAGC 360
GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCAGGA GTACCGCATC CAGCTCCGCG 420
75 TCTACAAAGC TCGGAGGAG CCAACATCC AGGTCAACCC CCTGGGCATC CCTGTGAACA 480
GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGTACCCC ATTCTCAAG 540
TCATCTGTGA CAAGAATGCG CGGCTCTGA AGGAGGAGAA GAACCGGCTC CACATTCAGT 600
CGTCCAGAGC TGTGAGGTGG AGTGTTTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC 660
TGGTTAAAGA AGACAAGAT GCCCAGTTT ACTGTGAGCT CAACCTACCG CTGCCAGTGC 720
80 GGAACCATG GAGGAGTCC AGGGAAGTCA CCGTCCCTGT TTTCTACCG ACAGAAAAAG 780
TGTGGCTGGA AGTGGAGCCC GTGGGAATGC TGAAGGAAGG GGACCGCGTG GAAATCAGGT 840
GTTTGGCTGA TGGCAACCTC CCACACACT TCAGCATCAG CAAGCAGAAC CCCAGCACCA 900
GGGAGGCAGA GGAAGAGACA ACCAAGGACA ACGGGGTCCT GGTGCTGGAG CCTGCCGGA 960
AGGAACACAG TGGCGCTAT GAAATGTCAG CCTGGAACCT GGACACCATG ATATCGCTGC 1020
85 TGAAGTGAAC ACAGGAAGT CTGGTGAAC ATGTGCTGTA CGTCCGAGT AGTCCCGCAG 1080
CCCTCAGAG ACAGGAAGGC AGCAGCCTCA CCCTGACCTG TGAGGCAGAG AGTAGCCAGG 1140
ACCTCAGATT CCAATGGCTG AGAGAAGAGA CAGACCAGGT GCTGGAAGAG GGGCCTGTGC 1200
```

	TTCACTTGCA	TGACCTGAAA	CGGGAGGAG	GAGGCGGCTA	TCGCTGCGTG	GGCTCTGTGC	1260
	CCAGCATACC	CGGCTTGAAC	CGCACACAGC	TGGTCAAGCT	GGCCATTTTT	GGCCCCCCTT	1320
	GGATGGCATT	CAAGGAGAGG	AAGGTGTGGG	TGAAAGAGAA	TATGGTGTG	AATCTGTCTT	1380
5	GTGAAGCGTC	AGGGCACCCC	CGGCCCCACA	TCTCCTGGAA	CGTCAACGGG	ACGGCAAGTG	1440
	AACAAGACCA	AGATCCACAG	CGAGTCTCTGA	GCACCTCGAA	TGTCTCGTG	ACCCCGGAGC	1500
	TGTTGAGAC	AGGTGTTGAA	TGCACGGCTT	CCAACGACCT	GGGCAAAAC	ACCAGCATCC	1560
	TCTTCTGGGA	GCTGGTCAAT	TTAACCACCC	TCACACCAGA	CTCCAACACA	ACCCTGGGCC	1620
	TCAGCACTTC	CATGTCACGT	CCTCATACCA	GAGCCAACAG	CACCTCCACA	GAGAGAAAGC	1680
10	TGCCGAGGCC	GGAGAGCCGG	GGCGTGGTCA	TCGTGGCTGT	GATTGTGTGC	ATCCTGGTCC	1740
	TGGCGGTGCT	GGGCGCTGTC	CTCTATTTC	TCTATAAGAA	GGGCAAGCTG	CCGTGCAGGC	1800
	GCTCAGGAA	GACGAGATC	ACGCTGCCCC	CGTCTCGTAA	GACCGAAGCT	GTAGTTGAAG	1860
	TTAAGTCAGA	TAAGCTCCCA	GAAGAGATGG	GGCTCCTGCA	GGGCAAGCAG	GGTGACAAGA	1920
	GGGCTCCGGG	AGACAGGGA	GAGAAATACA	TCGATCTGAG	GCATTAGCCC	CGAATCACTT	1980
	CAGCTCCCTT	CCCTGCCCTG	ACCATTCCTA	GCTCCCTGCT	CACCTCTTCT	TCAGCCAAAG	2040
15	CCTCCAAAGG	GACTAGAGAG	AAGCCTCCTG	CTCCCTCAC	CTGCACACCC	CCTTTCAGAG	2100
	GGCCACTGGG	TTAGGACCTG	AGGACCTCAC	TTGGCCCTGC	AAGCCGCTTT	TCAGGAGCCA	2160
	GTCCACCACC	ATCTCTTCCA	CGTTGAGTGA	AGCTCATCCC	AAGCAAGGAG	CCCCAGTCTC	2220
	CCGAGCGGGT	ASGAGAGTTT	CTTGACAGAC	GTGTTTTTTC	TTTACACACA	TTATGGCTGT	2280
20	AAATACCTCG	CTCCTGCCAG	CAGCTGAGCT	GGGTAGCCTC	TCTGAGCTGG	TTTCTGCTCC	2340
	CAAAGGCTGG	CTTCCACCAT	CCAGGTGCAC	CACCTGAAGT	AGGACACACC	GGAGCCAGGC	2400
	GCCTGCTCAT	GTGGAAGTGC	GCTGTTTACA	CCCGCTCCGG	AGAGCACCCC	AGCGCATATC	2460
	AGAAGCAGCT	GCAGTGTGTC	TGCCACCACC	CTCCTGCTCG	CCTCTTCAAA	GTCTCCTGTG	2520
	ACATTTTTTC	TTTGGT CAGA	ASCCAGGAAC	TGGTGTCAAT	CCTTAAAGAA	TACGTGCCGG	2580
	GGCCAGTACT	GGTGGCTCAG	GCCTGTAATC	CCAGCACTTT	GGGAGGCCGA	GGCGGGCGGA	2640
25	TCACAAAGTC	AGGACGAGAC	CATCCTGGCT	AACACGGTGA	AACCTGTCTT	CTACTAAAAA	2700
	TACAAAAAAA	AATTAGCTAG	GCGTAGTGGT	TGGCACCTAT	AGTCCCAGCT	ACTCGGAAGG	2760
	CTGAAGCAGG	AGAAATGGTAT	GAATCCAGGA	GGTGAGCTT	GCAGTGAGCC	GAGACCGTGC	2820
	CACGTGACTC	CAGCTCGGGC	AACACAGCGA	GACTCCGTCT	CGAGGAAAAA	AAAAAGAAAG	2880
30	ACGGGTACCT	CGGTGAGGAA	AGCTGGGGGC	TGTTTTGAG	TTGAGGTGAA	TTAGCCTCAA	2940
	TCCCGGTGTT	CACCTGTCTC	CATAGCCCTC	TTGATGATC	ACGTAAAACT	GAAAGGCAGC	3000
	GGGAGCAGA	CAAGATAGAG	GTCTACACTG	TCCTTCATGG	GGATTAAAGC	TATGGTTATA	3060
	TTAGCACCAA	ACTTCTACAA	ACCAAGCTCA	GGGCCCCAAC	CCTAGAAGGG	CCCAATGAG	3120
	AGAATGTATC	TTAGGGATGG	AAAACGGGGC	CTGGCTAGAG	CTTGGGTGTG	GTGTGTCTGT	3180
35	CTGTGTGTAT	GCATACATAT	GTGTGTATAT	ATGGTTTTGT	CAGGTGTGTA	AATTTGCAAA	3240
	TTGTTTCTTT	TATATATGTA	TGTATATATA	TATATGAAAA	TATATATATA	TATGAAAAAT	3300
	AAAGCTTAAT	TGTCCAGAAA	AATCATACAT	TGCTTTTTTA	TTCTACATGG	GTACCAACAG	3360
	AACCTGGGGG	CCTGTGAAAC	TACAACCAAA	AGGCACACAA	AACCGTTTCC	AGTTGGCAGC	3420
	AGAGATCAGG	GGTTACCTCT	GCTTCTGAGC	AAATGGCTCA	AGCTCTACCA	GAGCAGACAG	3480
40	CTACCTACTT	TTTCCAGCAG	AAAACGTCCC	GTATGACGCA	GCACGAAGGG	CCTGGCAGGC	3540
	TGTTAGCAGG	AGCTATGTCC	CTTCTATCG	TTTCCGTCCA	CTT		

Seq ID NO: 553 Protein sequence
Protein Accession #: NP_006491.1

45	1	11	21	31	41	51	
	GLPLRLVCAFL	LAACCCCPRV	AGVPGAEAPQ	APELVEVEVG	STALLKCGLS	QSQGNLSHVD	60
	WFSVHKEKRT	LIFRVRGQGG	QSEPGVEYQR	LSLQDRGATL	ALTQVTPQDE	RIFLCQGRKP	120
50	RSQERYRIQLR	VYKAPPEEPNI	QVNPLGIPVN	SKEPEEVATC	VGRNGYPIPO	VIWYKNGRPL	180
	KEEKNRVHIQ	SSQTVESGSL	YTLQSLKQAQ	LKVEDKDAQF	YCELNYRLPS	GNHMKESREV	240
	TVPVFFYPTK	VHLEVEPVGM	LKEDRVEIR	CLADGNPPPH	FSISKQNPST	REAEETND	300
	NGVLVLEPAR	KEHSGRYEQC	AWNLDTMISL	LSEPEQLLVN	YVSDVRVSPA	APERQEGSSL	360
	TLTCEAESQ	DLEPQWLRRE	TQVLERGFV	LQLHDLKREA	GGGYRCVASV	PSIPGLNRTQ	420
55	LVLKLAIFGPP	WMAPKERVN	VKENMVLNLS	CEASGHPRPT	ISWNVNGTAS	EQDQDPQRLV	480
	STLANVLVPE	LLETGVECTA	SNDLGKNTSI	LFLELVNLTT	LTPDSNTTIG	LSTSTASPHET	540
	RANSTSTERK	LFEPBSRGVV	IVAVIVCILV	LAVLGAULVF	LYKRGKLPKR	RSQKEITLTP	600
	PSRKTELVEE	VKSKLPEEM	GLLQSSQDK	RAPGDQGEKY	IDLRH		

Seq ID NO: 554 DNA sequence
Nucleic Acid Accession #: NM_003183.3
Coding sequence: 165..2639

65	1	11	21	31	41	51	
	TCGAGCCTGG	CGGTAGAATC	TTCCCACTAG	GCGGCGCGGG	AGGGAAAAGA	GGATTGAGGG	60
	GCTAGGCGGG	CGCGATCCCG	TCCTCCCCCG	ATGTGAGCAG	TTTTCGAAA	CCCCGTGAGG	120
	CGAAGGCTGC	CCAGAGAGGT	GGAGTCGGTA	GCGGGGCGGG	GAACATGAGG	CAGTCTCTCC	180
	TATTCCTGAC	CAGCGTGGTT	CCTTCTGTGC	TGGCGCGCGG	ACCTCCGGAT	GACCCGGGCT	240
70	TCGGCCCCCA	CCAGAGACTC	GAGAAGCTTG	ATTCTTTGCT	CTCAGACTAC	GATATTCTCT	300
	CTTTATCTAA	TATCCAGCAG	CATTGCGTAA	GAAAAAGAGA	TCTACAGACT	TCAACACATG	360
	TAGAAACACT	ACTAACTTTT	TCAGCTTTGA	AAAGGCATTT	TAAATTATAC	CTGACATCAA	420
	GTACTGAACG	TTTTTCAAAA	AATTTCAAGG	TCGTGGTGGT	GGATGGTAAA	AACGAAAGCG	480
	AGTACACTGC	AAAATGGCAG	GACTTCTTCA	CTGGACACGT	GGTTGGTGAG	CCTGACTCTA	540
75	GGGTTCTAGC	CCACATAAGA	GATGATGATG	TTATAATCAG	AATCAACACA	GATGGGGCCG	600
	AATATAACAT	AGAGCCACTT	TGGAGATTGG	TTAATGATAC	CAAAGACAAA	AGAATGTTAG	660
	TTTATAAACT	TGAAGATATC	AAGAATGTTT	CACGTTTGCA	GTCTCCAAAA	GTGTGTGGTT	720
	ATTTAAAGT	GGATAATGAA	GAGTTGCTCC	CAAAAGGGTT	AGTAGACAGA	GAACCACTG	780
	AAGAGCTTGT	TCATCGAGTG	AAAAGAAGAG	CTGACCCAGA	TCCATGAAG	AACAGTGTA	840
80	AATTATTGGT	GGTAGCAGAT	CATCGCTTCT	ACAGATACAT	GGGCAAGGGG	GAAGAGAGTA	900
	CAACTACAAA	TACTTAATA	GAGCTAATTG	ACAGAGTTGA	TGACATCTAT	CGGAACACTT	960
	CATGGGATAA	TGCAAGTTTT	AAAGGCTATG	GAATACAGAT	AGAGCAGATT	CGCATTTCTA	1020
	AGTCTCCACA	AGAGGTAAAA	CCTGGTGAAG	AGCACTACAA	CATGGCAAAA	AGTTACCCAA	1080
	ATGAAGAAAA	GGATGCTTGG	GATGTGAAGA	TGTTGTGAGA	GCAATTTAGC	TTTGATATAG	1140
85	CTGAGGAAGC	ATCTAAAGTT	TGCTTGGCAC	ACCTTTTTCAC	ATACCAAGAT	TTTGATATGG	1200
	GAACCTCTTG	ATTAGCTTAT	GTGGCTCTC	CCAGAGCAAA	CAGCCATGGA	GGTGTGTTGC	1260
	CAAAGGCTTA	TTATAGCCCA	GTTGGGAAGA	AAAATATCTA	TTTGAATAGT	GGTTTGAAGA	1320
	GCACAAAGAA	TTATGGTAAA	ACCATCTTAA	CAAAGGAAGC	TGACCTGGTT	ACAATCATG	1380

	AATTGGGACA	TAATTTTGGG	GCAGAACATG	ATCCGGATGG	TCTAGCAGAA	TGTGCCCCGA	1440
	ATGAGGACCA	GGGAGGGAAA	TATGTCAATG	ATCCCATAGC	TGTGAGTGGC	GATCAGGAGA	1500
	ACATAAGATG	GTTTTCAAAC	TGCAGTAAAC	AATCAATCTA	TAAGACCAAT	GAAGTAAAGG	1560
5	CCGAGAGTGG	TTTTCAAGAA	CGCAGCAATA	AAGTTTGTGG	GAACTCGAGG	GTGGATGAAG	1620
	GAGAAGAGTG	TGATCCTGGC	ATCATGTATC	TGAACAACGA	CACCTGCTGC	AACAGCGACT	1680
	GCACGTTGAA	GGAGAGGTGC	CAGTGCAGTG	ACAGGAACAG	TCCTTGCTGT	AAAACTGTCT	1740
	AGTTTGAGAC	TGCCAGGAAG	AAGTGCCAGG	AGGCGATTAA	TGCTACTTGC	AAAGCGGTGT	1800
	CCTACTGCAC	AGGTAAATAGC	AGTGAGTGCC	CGCCTCCAGG	AAATGCTGAA	AATGACACTG	1860
	TTTGCTTTGA	TCTTGGCAAG	TGTAAGGATG	GGAAATGCAT	CCCTTTCTGC	GAGAGGGAAC	1920
10	AGCAGCTGGA	GTCTCTGTGA	TGTAATGAAA	CTGACAACCT	CTGCAAGGTG	TGCTGCAGGG	1980
	ACCTTTCTGG	CCGCTGTGTG	CCCTATGTGC	ATGCTGAACA	AAAGAACTTA	TTTTTGAGGA	2040
	AAGGAAAGCC	CTGTACAGTA	GGATTTTGTG	ACATGAATGG	CAAAATGTAG	AAACGAGTAC	2100
	AGGATGTAAT	TGAACGATTT	TGGGATTTC	TTGACCAGCT	GAGCATCAAT	ACTTTTGAAA	2160
	AGTTTTTAGC	AGACAACATC	GTTGGGTCTG	TCCTGGTTTT	CTCCTTGATA	TTTTGGATTG	2220
15	CTTTCAGCAT	TCTTGTCCAT	TGTGTGGATA	AGAAATTTGA	TAAACAGTAT	GAATCTCTGT	2280
	CTCTGTTC	CCCCAGTAA	GTGAAATG	TGAGCAGCAT	GGATTCGCA	TCGGTTCGCA	2340
	TTATCAAAAC	CTTCTGTG	CCCCAGACT	CAGGCGCCT	GCAGCCTGCC	CCTGTGATCC	2400
	CTTCGGCGCC	AGCAGCTCCA	AACTGGAGC	ACCAGAGAAT	GGACACCATC	CAGGAAGACC	2460
20	CCAGCAGAGA	CTCCCATATG	GACGAGGATG	GGTTTGAGAA	GGACCCCTTC	CCAAATAGCA	2520
	GCACAGCTGC	CAAGTCAATT	GAGGATCTCA	CGGACCATCC	GSTCGCCAGA	AGTGAAGAGG	2580
	CTGCTCTCTT	TAAACTGCG	CGTCAGAAAT	GTGTTAACAG	CAAGAAACAA	GAGTGCTAAT	2640
	TTAGTTCTCA	GCTCTTCTGA	CTTAAGTGTG	CAAAATATTT	TTATAGATTT	GACCTACAAA	2700
	TCAATCACAG	CTTGATTTTT	GTGAAGACTG	GGAAAGTACT	TAGCAGATGC	TGGTCAATGT	2760
	TTTGAACCTC	CTGCAGGTAA	ACAATTCTTG	TGTGGTTTGG	CCCTTCTCCT	TTTGAAGAGG	2820
25	TAAAGTGAAA	GTGAATCTAC	TTATTTTGTG	GCTTTCAGGT	TTTAGTTTTT	AAAATATCTT	2880
	TTGACCTGTG	GTGCAAAAGC	AGAAAAATCA	GCTGGATTGG	GTATGAATA	TTTACGTTTT	2940
	TGTAATTTAA	TCTTCTATAT	TGATAACAGC	ACTGACTAGG	GAAATGATCA	GTTTTTTTTT	3000
	ATACACTGTA	ATGAACCGCT	GAATATGAAG	CATTTGGCAT	TTATTTGTGA	GAAGAGTGGA	3060
	ATAGTTTTTT	TTTTTTTTTT	TTTTTTTTTG	CTTCAACTAA	AAACAAAGGA	GATAAATTTA	3120
30	GTATACATGT	TATCTAAAT	GTGGTCTAT	TCTAGTTAT	TACCCAGAGT	TTTTATGTAG	3180
	CAGGAAATAT	ATATCTCTAA	ATTTAGAAAT	CATTTGGGTT	AATATGGCTC	TTCTAATTC	3240
	TAAGACTAAT	GCTCAGAAAC	TAACCACTAC	CTTACAGTGA	GGGCTATACA	TGGTAGCCAG	3300
	TTGAATTTAT	GGAACTTACC	AACGTGTTAG	GGCCCTGATT	TGCTGGGCAG	TTTTTCTGTA	3360
35	TTTTATAAGT	ATCTTCATGT	ATCCCTGTTA	CTGATAGGGA	TACATGCTTT	AGAAAATTTA	3420
	CTATTGGCTG	GGAGTGGTGG	CTCATGCTGT	TAATCCCGAG	ACTTGGAGAG	GCTGAGGTTG	3480
	CGCCACTACA	CTCCAGCCTG	GGTGACAGAG	TGAGATCTGC	CTC		

Seq ID NO: 555 Protein sequence
Protein Accession #: NP_003174.2

40	1	11	21	31	41	51	
	MRQSLFLTS	VVPFVLAPRP	PDDPGFPGHQ	RLEKLDLSLS	DYDILSLNSI	QQHSVRKRDL	60
	QTSHEVELL	TFSALKRHKF	LYLTSSTERF	SNFKVUVVD	GNSESYTAK	WQDFPTGHVV	120
45	GEPSRVLAE	IRDDVLIIR	MTDGAENYIE	PLWRPVNDTK	DKRMLVYKSE	DIRNVSLQS	180
	PKVCGYLKVD	NEELLPKGLV	DREPEELVH	RVKRRADPDP	MONTCKLLVV	ADHRFYRYMG	240
	RGEESTTNY	LIELIDRVDD	TYRNTSWDNA	GFKGYGQIE	QIRILKSPQE	VKPKGEKHYNM	300
	AKSYPNEEKD	AWDKVMLLEQ	PSFDIAEAS	KVCLAHLFY	QDFDMGTGL	AYVGSPPRNS	360
	HGGVCPKAYI	SPVKGKNIYL	NSGLTSTKRY	GRTILTKEAD	LVTTHELGHV	FGAEHDPDGL	420
50	AECAPNEDQG	GKVMYPIAV	SGDHENKMF	SNCSKQSIYK	TIESKAQECF	QERSNKVCGN	480
	SRVDEGEBCD	PGIMYLNMDT	CCNSDCTLKE	GVQCSDRNSP	CCKNQCQFETA	QKKCQEAALNA	540
	TKGVSYSYCT	NSSECPPEPN	AENDTVCLDL	GKCKDGKCI	PCEREQQLS	CACNETDNCS	600
	KVCCRDLDSR	CVPVYDAEQK	NLFLRKKGKPC	TVGFCDMNGK	CEKRVQDVIE	RFWDFIDQLS	660
	INTFGKFLAD	NIVGSLVLFQ	LIFWIPFSIL	VHCVDKLDK	QYESLSLFLP	SNVEMLSMD	720
55	SASVRIIKFP	PAPQTPGRLO	PAPVIPSAPA	APKLDHQRM	TIQEDPSTDS	HMEDEGFEKD	780
	PPNPSSTAAR	SPEDLTDFHV	ARSEKAASF	LQRQNRVNSK	BTEC		

Seq ID NO: 556 DNA sequence
Nucleic Acid Accession #: NM_021832.1
Coding sequence: 164..2248

60	1	11	21	31	41	51	
	TOGAGCTGG	CGGTAGAATC	TTCCAGTAG	CGGCGCGGG	AGGAAAAGAG	GATTGAGGGG	60
65	CTAGGCGGG	CGGATCCCGT	CCTCCCCGA	TGTGAGCAGT	TTTCCGAAAC	CCCGTCAGGC	120
	GAAGGCTGCC	CAGAGAGGTG	GAGTCGGTAG	CGGCGCGGG	AACATGAGGC	AGTCTCTCCT	180
	ATTCTGACC	AGCGTGGTTC	CTTCTGTGCT	GGCGCCGCGA	CCTCCGATG	ACCCGGGCTT	240
	CGGCCCCAC	CAGAGACTCG	AGAAGCTTGA	TTCTTTGCTC	TCAGACTAAG	ATATTCTCTC	300
70	TTTATCTAAT	ATCCAGCAGC	ATTCGGTAAG	AAAAAGAGAT	CTACAGACTT	CAACACATGT	360
	AGAAACACTA	CTAACTTTTT	CAGCTTTGAA	AAGGCATTTT	AAATTATACC	TGACATCAAG	420
	TACTGAACGT	TTTTTCAAAA	ATTTCAAGGT	CGTGGTGGTG	GATGTAATA	ACGAAAGCGA	480
	GTACACTGTA	AAATGGCAGG	ACTTCTTCAC	TGACACAGTG	GTGTTGAGC	CTGACTCTAG	540
	GGTTCTAGCC	CACATAAGAG	ATGATGATGT	TATAATCAGA	ATCAACACAG	ATGGGGCCGA	600
75	ATATAACATA	GAGCCACTTT	GGAGATTGT	TAATGATACC	AAAGACAAAA	GAATGTTAGT	660
	TTATAAATCT	GAAGATATCA	AGAAATGTTT	ACGTTTGACG	TCTCCAAAAG	TGTGTGGTTA	720
	TTTAAAGTG	GATAATGAAG	AGTTGCTCCC	AAAAGGGTTA	GTAGACAGAG	AACCACTGA	780
	AGAGCTTGTT	CATCGAGTGA	AAAGAAGAGC	TGACCCAGAT	CCCATGAAGA	ACACGTGTAA	840
	ATTATTGGTG	GTAGCAGATC	ATCGCTTCTA	CAGATACATG	GGCAGAGGGG	AAGAGAGTAC	900
80	AACACAAAT	TACTTAATAG	AGCTAATTGA	CAGAGTTGAT	GACATCTATC	GGAAACATTC	960
	ATGGGATAAT	GCAGGTTTTA	AAGGCTATGG	AATACAGATA	GAGCAGATTC	GCAATCTCAA	1020
	GTCTCCACAA	GAGGTAAAAC	CTGGTGAATA	GCATACAAAC	ATGGCAAAA	GTTACCCAAA	1080
	TGAAGAAAAG	GATGCTTGGG	ATGTGAAGAT	GTGCTAGAG	CAATTTAGCT	TTGATATAGC	1140
	TGAGGAAGCA	TCTAAAGTTT	GCTTGGCACA	CCTTTTCACA	TACCAAGATT	TTGATATGGG	1200
85	AACCTTTGA	TTAGCTTATG	TTGGCTCTCC	CAGAGCAAAC	AGCCATGGAG	GTGTTTGTCC	1260
	AAAGGCTTAT	TATAGCCCAT	TTGGGAAGAA	AAATATCTAT	TTGAATAGTG	GTGAGCAGAG	1320
	CACAAAGAAT	TATGGTAAAA	CCATCCTTAC	AAAGGAAGCT	GACCTGGTTA	CAACTCATGA	1380
	ATTGGGACAT	AATTTTGGAG	CAGAACATGA	TCCGATGGT	CTAGCAGAA	GTGCCCCGAA	1440

5
10
15
20
25
30
35

TGAGGACCAG GGAGGGAAAT ATGTCATGTA TCCCATAGCT GTGAGTGGCG ATCACGAGAA 1500
CAATAAGATG TTTTCAAACT GCAGTAACA ATCAATCTAT AAGACCATTG AAAGTAAGGC 1560
CCAGGAGTGT TTTCAAGAAC GCAGCAATAA AGTTTGTGGG AACTCGAGGG TGGATGAAGG 1620
AGAAAGATGT GATCCTGGCA TCATGTATCT GAACAACGAC ACCTGCTGCA ACAGGAGCTG 1680
CACGTTGAAG GAAGGTGTCC AGTGCACTGA CAGGAACAGT CCTTGCTGTA AAAACTGTCA 1740
GTTTGAGACT GCCCAGAAGA AGTGCCAGGA GCGGATTAAT GCTACTTGCA AAGGCGTGTC 1800
CTACTGCACA GGTAAATAGCA GTGAGTGCCC GCCTCCAGGA AATGCTGAAG ATGACACTGT 1860
TTGCTTGGAT CTGCGCAAGT GTAAGGATGG GAAATGCATC CCTTCTGCG AGAGGGAACA 1920
GCAGCTGGAG TCCTGTGCAT GTAATGAAAC TGACAACTCC TGCAAGGTGT GCTGCAGGGA 1980
CCTTCCGGC CGCTGTGTGC CCTATGTGTA TGCTGAACAA AAGAACTTAT TTTTGAGGAA 2040
AGGAAAGCCC TGTACAGTAG GATTTTGTA CATGAATGCG AAATGTGAGA AACGAGTACA 2100
GGATGTAAT GAACGATTTT GGGATTTTAT TGACCACTG AGCATCAATA CTTTGGGAAA 2160
GTTTGTAGCA GACAACATCG TTGGGTCTGT CTGCTTTC TCCTTGATAT TTTGGAATCC 2220
TTTCAGCAT CTGTGTGCAT GTGTGTAACG TCGAAATGCT GAGCAGCATG GATTCTGCAT 2280
CGGTTGCGAT TATCAAAACC TTTCTGCGC CCCAGACTCC AGGCCGCTG CAGCCTGCCC 2340
CTGTGATCCC TTGCGGCCCA GCAGCTCCAA AACTGGAGCA CCAGAGAATG GACACCATCC 2400
AGGAAGACCC CAGCAGACAG TCACATATGG ACGAGGATGG GTTTGAGAAG GACCCCTTCC 2460
CAATAGCAG CACAGCTGCC AAGTCATTG AGGATCTCAC GGACCATCG GTCCACCAGAA 2520
GTGAAAAGGC TGCTCTCTT AAACCTGCAG CTCAGAATCG TGTGACAGC AAAGAAACAG 2580
AGTGCTAAT TAGTCTCAG CTCCTCTGAC TTAAGTGTGC AAAATATTTT TATAGATTG 2640
ACCTACAAT AATCACAGT TATATTTTGT GAAGACTGGG AAGTGACTTA CGAGATGCTG 2700
GTACGTGTT TGAACCTCT GCAGGTAAAC AGTTCTGTG TGGTTTGGCC CTTCTCTTT 2760
TGAAAAGGTA AGGTGAAGGT GAATCTAGCT TATTTTGAGG CTTTCAGGTT TTAGTTTITA 2820
AAATATCTT TGACCTGTG TGCAAAAGCA GAAATACAG CTGGATTGGG TTATGAGTAT 2880
TTACGTTTT GTAAATTAAT CTTTATATAT GATAACAGGC ACTGACTAGG GAAATGATCA 2940
GTTTTTTTT ATACACTGTA ATGAACCGCT GAATATGAAG CATTTGGCAT TTATTTGTGA 3000
GAAAAGTGA ATAGTTTTTT TTTTTTTTT TTTTTTTTGC CTTCAACTAA AAACAAAGGA 3060
GATAAATTA GTATACATTG TATCTAAAT GTGGGTCTAT TTCTAGTTAT TACCCAGAGT 3120
TTTTATGTAG CAGGGAATAA ATATATCTAA ATTTAGAAAT CATTTGGGTT AATATGSCCTC 3180
TTCATAATC TAAGACTAAT GCTCAGAAC TAACCACTAC CTTACAGTGA GGGCTATACA 3240
TGGTAGCCAG TTGAATTTAT GGAATCTACC AACTGTTAG GGCCTGATT TGCTGGGCAG 3300
TTTTCTGTA TTTTATAAG ATCTTCATGT ATCCCTGTTA CTGATAGGGA TACATGTCTT 3360
AGAAAATTCA CTATTGGCTG GAGTGTGTG CTCATGCCG TAATCCAGC ACTTGGAGAG 3420
3421 GCTGAGGTTG CGCCACTACA CTCAGCCTG GGTGACAGAG TGAGATCTGC CTC

Seq ID NO: 557 Protein sequence
Protein Accession #: NP_068604.1

40
45
50

1 11 21 31 41 51
MRQSLFLTS VVPFVLAPRP PDDPGFGPHQ RLEKLDLSLS DYDILSLNSI QQHSVRKRD 60
QTSTHVELL TFSALKRHF LVLTSSTERF SQNFVUVVD GKNESEYTVK WQDFPTGHV 120
GEPDSRVLAH IRDDVVIIRI NTDGAENIE PLWRPVNDTK DKRMLVYKSE DIKNVSRLOS 180
PKVCGYLKVD NEELLPLGLV DREPPPELVH RVKRRADPDP MKNTCKLLVV ADHRFPYRYMG 240
RGEESTTNY LIEILDRVDD IYRNTSWDNA GFKGYGQIE QIRILKSPQE VKPGEKHYNM 300
AKSYFNEEKD AWDVKMLLEQ FSEFDIAEAS KVCLAHLFTY QDFDMGTGLL AYVGSFRANS 360
HGGVCPKAYY SPVKGKNIYL NSGLTSTKNY KRTILTKEAD LVTTHELGHN FGAERDPPDL 420
AECAPNEDQG GKYVMYPIAV SGDHENNKMF SNCSKQSIYK TIESKAQECF QERSNKVCGN 480
SRVDEGEEDQ PGIMYLNNDT CCNSDCTLKE GVQCSDRNSP CCKNCQFETA QKKCQEAINA 540
TCKGVSYCTG NSSBCPPPGN AEDDTVCLDL GKCKDGKICP FCEREQQLS CACNETDNCS 600
KVCCRDLSGR CVFYVDAEQK NLFLRKGKPC TVGFCDMNGK CEKRVQDVIE RFWDFIDQLS 660
INTFGKPLAD NIVGSVLVPS LIFWIPPSIL VHCV

Seq ID NO: 558 DNA sequence
Nucleic Acid Accession #: NM_004994.1
Coding sequence: 20..2143

60
65
70
75
80
85

1 11 21 31 41 51
AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCTCG GTCTGCTGTC TCCTGGTGCT 60
GGGCTGCTGC TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTGTGCTCT TCCTGGAGA 120
CCTGAGAACC AATCTCACCG ACAGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGITA 180
CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGGCG TGCTGCTTCT 240
CCAGAGCAA CTGCTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300
GCGAACCCCA CGGTGCGGGG TCCAGACCTT GGCAGATTTC CAAACCTTTG AGGGCGACCT 360
CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGAAG ACTTGCCGCG 420
GGCGGTGATT GACGACGCTT TTGCCCGCGC CTTCGCACTG TGAGCGCGCG TGACGCCGCT 480
CACCTTCACT CGCGTGTACA GCGGGGAGCG AGACATCGTC ATCCAGTTTG GTGTGCGGGA 540
GCACGGAGAC GGGTATCCCT TGACGGGAA GGACGGGCTC CTGGCACAGC CCTTCTCTCC 600
TGGCCCGGCG ATTCAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCTTGGGCAA 660
GGGCGTCTGT GTTCAACTC GGTTTGGAAG CGCAGATGGC GCGGCTGCC ACTTCCCTCT 720
CATCTTCGAG GGCCTCTCT ACTCTGCTG CACCACCGAC GGTGCTCCG ACGCTTGGC 780
CTGGTGCAGT ACCACGGCCA ACTAOGACAC CGACGACCGG TTTGGCTTCT GCCCAGCGA 840
GAGACTCTAC ACCCGGGAGA GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT 900
CCAAGGCCAA TCCTACTCG CCTGCACCAC GGCAGGTGCG TCCAGCGGCT ACCGCTGGTG 960
CGCCACCAAC GCCAATACG ACGGGGACAA GCTCTTGGCG TTCTGCCGA CCGAGCTGA 1020
CTCGAGGTG ATGGGGGGCA ACTCGGGCGG GGAGCTGTGC GTCTTCCCTT TCACCTTCTT 1080
GGTAAGGAG TACTCGACCT GTACAGCGA GGGCCGCGGA GATGGGCGCC TCTGTGCGC 1140
TACCACCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200
TTTGTCTCTC GTGGCGGCGC ATGAGTTGCG CCACGCGCTG GGCTTAGATC ATTCTCTAGT 1260
GCCGGAGCGC CTATGTATCC CTATGTACCG CTTCCTGAG GGGCCCCCTT TGCATAAGGA 1320
CGAGTGAAT GGCATCGGCG ACCTCTATGG TCCTCGCCCT GAACCTGAGC CACGCGCTCC 1380
AACCACCAAC ACCCGGAGC CACCGGCTCC CCGACGGTCT TGCCCCACCG GACCCCTCAC 1440
TGTCACCCCC TCAGAGCGCC CACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC 1500
AGTCCCCCCC ACTGCTGGCC CTTCTAOGGC CACTACTGTG CCTTTGAGTC CGGTGAGGA 1560
TGCTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAAACAGC TGTATTTGTT 1620
CAAGGATGGG AAGTACTGCG GATTCTCTGA GGGCAGGGGG AGCCGCGCGC AGGGCCCCCTT 1680

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

CCTATGCGC GACAAGTGGC CGCGCTGCC CCGCAAGCTG GACTCGGTCT TTGAGGAGCC 1740
 GCTCTCCAAG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC 1800
 GGTGCTGGGC CGAGGGCGTC TGGACAAGCT GGGCCCTGGG GCGGAGCTGG CCCAGGTGAC 1860
 CGGGGCCCTC CGAGTGGCA GGGGAAGAT GCTGCTGTTC AGCGGGCGGC GCGCTCTGGAG 1920
 GTTCGAGCTG AAGCGCGAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT 1980
 CCGCGGGGTG CCTTTGGACA CGCACGACGT CTTCAGTAC CGAGAGAAAG CCTATTCTGT 2040
 CCAGGAGCGC TTCTACTGGC GCGTGAGTTC CCGGAGTGAG TTGAACCAAG TGGACCAAGT 2100
 GGGCTACGTG ACCTATGACA TCCTCGAGTG CCTCGAGGAC TAGGGCTCCC GTCTCTCTTT 2160
 GCACTGGCAT GTAAATCCCT ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCCGGATA 2220
 CAAACTGGTA TTCTGTCTGT GAGGAAAGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC 2280
 TCACCTTTGT TTTTGTGTGG AGTGTCTCTA ATAAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 559 Protein sequence
 Protein Accession #: NP_004985.1

1 11 21 31 41 51
 MSLWQPLVLV LVLVGGCFPA PRQRQSTLVL FPGDLRLTNT DRQLAEYLY RYGYTRVAEM 60
 RGEKSLGPA LLLKQLSL PETGELDSAT LKAMRTPROG VPDILGRFQTF EBDLKHWHEN 120
 IYWIQNYSE DLPRAVIDDA FARAFALWSA VTPLTFTRVY SRDADIVIQF GVAEHGDGYP 180
 PDGKDLGALH AFPPGPGIQG DAHPDDDELW SLKGQVVVPT RFGNADGAAC HFFFIPEGRS 240
 YSACTTDSRS DGLPWCSTTA NYDTDDRFGF CPSERLYTRD GNADGKPCQF PFIFQGQSYS 300
 ACTTDDGRSD YRWCAATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTLFGKEYST 360
 CTSEGRGDGR LMCATTENFD SDKKWGFPCD QGYSLFLVAA HEFHALGLD HSSVPEALMY 420
 PMYRFTGPP LHKDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAPPTVCPT GPPTVHPSE 480
 PTAGTGPPTS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDLALAEIGNQ LYLFDKGYW 540
 RFSBGRGSRP GPFLLIADKW PALPRKLDV FEEPLSKLFP FFSGRQVVVY TGASVLGPRR 600
 LDKLGLGADV AQVTGALRSR RGMILLFSGR RLWRFDVKAQ MVDPRSAEV DRMPGVPDL 660
 THDVQYREK AYPCQDRFYW RVSSRSELNQ VDQGVYVTD ILQCPED

Seq ID NO: 560 DNA sequence
 Nucleic Acid Accession #: NM_000213.1
 Coding sequence: 127..5385

1 11 21 31 41 51
 CGCCCGCGCG CTGCAGCCCC ATCTCCTAGC GGCAGCCAG GCGCGGAGGG AGCAGTCCG 60
 CCCGAGGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GGCTGGCCGG GAGAGGGAGG 120
 AAGAGGATGG CAGGCGCCAG CCCAGGCCCA TGGGCCAGGC TGCTCCTGGC AGCCTTGATC 180
 AGCGTCAGCG TCTCTGGGAC CTCTGCAAAAC CGCTGCAAGA AGGCCCAAGT GAAGAGCTGC 240
 ACGGAGTGTG TCCGTGTGGA TAAAGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC 300
 CGGCGCTGCA ACACCCAGGC GGAGCTGCTG GCGCGGGCT GCCAGCGGGA GAGCATCGTG 360
 GTCATGAGGA CGAGCTTCCA AATCACAGAG GAGACCCAGA TTGACACCAC CCTGCGGCGC 420
 AGCCAGATGT CCCCCAAGG CCTGCGGGTC CGTCTGCGGC CCGGTGAGGA GCGGCATTTT 480
 GAGCTGAGCG TGTCTGAGCC ACTGGAGAGC CCGGTGGAAC TGTACATCCT CATGGACTTC 540
 TCCAACTCCA TGTCCGATGA TCTGGACAAC CTCAAGAGA TGGGGCAGAA CCTGGCTCGG 600
 GTCTGAGGCC AGCTACACAG GACTACACT ATTGSAATTG GCAAGTTTGT GACACAAGTC 660
 AGCGTCCCGC AGACGAGCAT GAGGCCTGAG AAGCTGAAGG AGCCCTGGCC CAACAGTGAC 720
 CCCCCCTTCT CCTTCAAGAA CGTCATCAGC CTGACAGAA ATGTGGATGA GTTCGGGAAT 780
 AAACGTCAGG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CTGAGGGCGG CTTGATGCC 840
 ATCTCTGACA CAGCTGTGTG CACGAGGGAC ATTGGCTGGC GCGCGGACAG CACCCACCTG 900
 CTGGTCTTCT CCACCGAGTC AGCCTTCCAC TATGAGGCTG ATGGGCGCAA CGTGTGGCT 960
 GGCATCATGA GCGCAACGA TGAACGCTGC CACCTGGACA CCACGGGCAC CTACACCCAG 1020
 TACAGGACAC AGGACTACCC GTCCGTGCCC ACCCTGGTGC GCGTCTCGC CAAGCACAAC 1080
 ATCATCCCCA TCCTTGTCTG CACCAACTAC TCCTATAGCT ACTACGAGAA GCTTCACACC 1140
 TATTTCCCTG TCTCTCACT GGGGGTGTG CAGGAGGACT CGTCAACAT CGTGAGCTG 1200
 CTGGAGGAGG CCTCAATCG GATCGCTCC AACCTGACA TCOCGCGCCT AGACAGCCCC 1260
 CGAGGCCCTT GGCAGAGGT CACCTCCAAG ATGTTCCAGA AGACGAGGAC TGGGTCTTTT 1320
 CACATCCGCG GGGGGGAAGT GGGTATATAC CAGGTGCAGC TGCGGGCCCT TGAGCACGTT 1380
 GATGGAGCGC ACGTGTGCCA GCTGCGGAG GAOCAGAGG GCAACATCCA TCTGAAACT 1440
 TCCTTCTCGG ACGGCTCAA GATGGAGCGG GGCATCATCT GTGATGTGTG CACCTGCGAG 1500
 CTGCAAAAG AGGTGCGGTC AGCTCGCTGC AGCTTCAACG GAGACTTCGT GTGCGGACAG 1560
 TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT 1620
 GACATTCAGC CCGCTCTGCG GAGAGGGGAG GACAAGCCGT GCTCCGCGCG TGGGGAGTGC 1680
 CAGTGGCGGC ACTGTGTGTG CTACGGGAGG GCGCGCTACG AGGGTCAAGT CTGCGAGTAT 1740
 GACAACCTCC AGTGTCCCG CACTTCCGGG TTCTCTGCA ATGACCGAGG ACGCTGCTCC 1800
 ATGGGCGCAG GTGTGTGTGA GCCTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCTCTCAG 1860
 AATGCCACTT GCATCGACAG CAATGGGGGC ATCTGTAAAT GACGTGGCCA CTGTGAGTGT 1920
 GGGCGCTGCC ACTGCCACCA GCAGTCCGTC TACACGACA CCATCTGCGA GATCAACTAC 1980
 TCGGCGATCC ACCCGGGCCT CTGCGAGGAC CTACGCTCCT GCGTGCAGTG CCAGGCGTGG 2040
 GGCACCGCGG AGAAGAAGGG GCGCACTGT GAGGAATGCA ACTTCAAGGT CAAGATGGTG 2100
 GACGAGCTTA AGAGAGCCGA GAGGTGGTG GTGCGCTGCT CTTCCCGGGA CGAGGATGAC 2160
 GACTGCACCT ACAGCTACAC CATGGAAGGT GACGGCGCCC CTGGGCCCCA CAGCACTGTC 2220
 CTGGTGACA AGAAGAAGGA CTGCCCCCG GCGTCTTCT GTTGGCTCAT CCCCCTGCTC 2280
 CTCTCTCTCC TGCGCTCTCT GCGCTGCTA CTGCTGCTAT GCTGGAAGTA CTGTGCTGTC 2340
 TGCAAGGCTC GCTTGCACCT TCTCCGCTGC TGCAACCGAG GTCAATGGT GGGCTTTAAG 2400
 GAAGACCACT ACATGCTGCG GGAGAACCTG ATGGCTCTGT ACCACTTGA CACGCCCATG 2460
 CTGCGCAGCG GGAACCTCAA GGGCGGTGAC GTGTGCTCGT GGAAGGTAC CAACCAATG 2520
 CAGCGCGCTG GCTTGGCCAC TCATGCCGCC AGCATCAACC CCACAGAGCT GGTGCCCTAC 2580
 GGGCTGTCTT TGCGCTGGC CCGCTTTG CCGAGAGAAC TGCTGAAGCC TGACACTCGG 2640
 GAGTGGCGCC AGCTGCCCA GAGGTGGAG GAGAACCTGA ACGAGGTCTA CAGGAGATC 2700
 TCGGTGTAC ACAAGCTCCA GCAGACCAAG TTCGCGCAGC AGCCCAATGC CGGGAAAAAG 2760
 CAAGACCCA CCAATTGTGA CACAGTGTG ATGGCGCCCC GCTCGGCCAA CGCGGCCCTG 2820
 CTGAAGCTTA CAGAGAAGCA GGTGGAACAG AGGGCTTCC ACACCTCAA GGTGGCCCCC 2880
 GGTCTACTCA CCTCTCTG AGACCAAGG GCGCGGGGCA TGGTGGAGTT CAGGAGGGC 2940
 GTGGAGCTGG TGAAGTACG GGTGCCCTC TTTATCCGC CTGAGGATGA CACAGAGAAG 3000
 CAGCTGCTGG TGGAGGCCAT CGAGTGCCC GCAGGCACCT CCACCTCTCG CCGCGGCTG 3060

	GTAAACATCA	CCATCATCAA	GGAGCAAGCC	AGAGACGTGG	TGTCCTTTGA	CGAGCCTGAG	3120
	TTCTCGGTCA	CCCGGGGGGA	CCAGGTGGCC	CGCATCCCTG	TCATCCGGCG	TGTCCTGGAC	3180
	GGCGGGAGT	CCCGAGTCTC	CTACCGCACA	CAGGATGGCA	CCGCGCAGGG	CAACCGGGAC	3240
5	TACATCCCGG	TGGAGGGTGA	GCTGCTGTTT	CAGCCTGGGG	AGGCTTGGAA	AGAGCTGCAG	3300
	GTGAAGCTCC	TGGAGCTGCA	AGAAGTTGAC	TCCCTCCTGC	GGGGCCGCCA	GGTCCGCGGT	3360
	TTCCAGCTCC	AGCTCAGCAA	CCCTAAGTTT	GGGGCCACC	TGGGCCAGCC	CCACTCCACC	3420
	ACCATCATCA	TCAGGGAGCC	AGATGAACCT	GACCGAGCT	TCACGAGTCA	GATGTTGTCA	3480
	TCACAGCCAC	CCCTCAGCG	CGACCTGGG	CGCCCGCAGA	ACCCCAATGC	TAAGCCCGCT	3540
	GGGTCCAGGA	AGATCCATT	CAACTGGCTG	CCCCCTTCTG	GCAAGCCAA	GGGGTACAGG	3600
10	GTAAAGTACT	GGATTACAGG	TGACTCCGAA	TCCGAAGCCC	ACCTGCTCGA	CAGCAAGGTG	3660
	CCCTCAGTGG	AGCTCACCAC	CCTGTACCCG	TATTGCGACT	ATGAGATGAA	GGTGTGCGCC	3720
	TACGGGGCTC	AGGGGAGGG	ACCCTACAGC	TCCCTGGTGT	CCTGCGCGAC	CCACCAGGAA	3780
	GTGCCAGCG	AGCCAGGGCG	TCTGGCCTTC	AATGTGCTCT	CCTCCACGGT	GACCCAGCTG	3840
	AGCTGGGCTG	AGCGGCTGA	GACCAACGGT	GAGATCACAG	CCTACGAGGT	CTGCTATGGC	3900
15	CTGGTCAACG	ATGACAACCG	ACCTATTGGG	CCCATGAAGA	AAGTGTGCTG	TGACAACCTT	3960
	AAGAACCAGG	TGCTGCTTAT	TGAGAACCTT	CGGGAGTCCC	AGCCCTACCG	CTACACGGTG	4020
	AAGGCGCGCA	ACGGGGCGCG	CTGGGGGCGT	GAGCGGGAGG	CCATCATCAA	CCTGGCCACC	4080
	CAGCCCAAGA	GGCCCATGTC	CATCCCATC	ATCCCTGACA	TCCCTATCGT	GGAGCGCCAG	4140
	AGCGGGGAGG	ACTACGACAG	CTTCTCTTATG	TACAGCGATG	ACGTTCTAAG	CTCTCCATCG	4200
20	GGCAGCCAGA	GGCCAGCGGT	CTCCGATGAC	ACTGAGCACC	TGGTGAATGG	CCGGATGGAC	4260
	TTTGCTCTCC	CGGGCAGCAC	CAACTCCCTG	CACAGGATGA	CCACGACCAG	TGCTGTCTGC	4320
	TATGGACATC	ACCTGACGCC	ACACGTGCCC	CACCGGTGCG	TAAGCACATC	CTCCACCTCT	4380
	ACACGGGACT	ACAACTCACT	GACCGCTCA	GAACTACTAC	ACTCGACCAC	ACTGCCGAGG	4440
25	GACTACTCCA	CCCTCACCTC	CGTCTCTCTC	CACGACTCTC	GCTGACTGTC	TGGTGTGCCC	4500
	GACACGCCCA	CCCGCTCTGT	GTTCTCTGCG	CTGGGGGCCA	CATCTCTCAG	AGTGAAGTGG	4560
	CAGGAGCGCG	GCTGCGAGCG	GCCGCTGCAG	GGCTACAGTG	TGGAGTACCA	GCTGCTGAAC	4620
	GGCGGTGATC	TGCATCGGCT	CAACATCCCC	AACCTGCCCC	AGACCTCGGT	GGTGGTGGAA	4680
	GACCTCCTCG	CCAAACACTC	CTACGTGTTT	CGCGTGCGGG	CCCAGAGCCA	GGAAAGGCTG	4740
30	GGCCGAGAGC	GTGAGGGTGT	CATCACCATT	GAATCCCAGG	TGCACCCGCA	GAGCCCACTG	4800
	TGTCCTCTGC	CTTCACTTTG	AGCACTCCCA	GTGCCCCAGG	CCCGCTGGTG	CCCGCTGGTG	4860
	TTCACTGCCC	TGAGCCGACA	CTCGCTGCAG	CTGAGCTGGG	AGCGGCCACG	GAGGCCCAAT	4920
	GGGATATATC	TCCGCTACCT	GGTGACCTGT	GAGATGGCCC	AAGGAGGAGG	GCCAGCCACC	4980
	GCATTCCGGG	TGGATGAGAG	CAGCCCGGAG	AGCCGGCTGA	CGGTGCGCGG	CCTCAGCGAG	5040
35	AACGTGCCCC	ACAAAGTTCAA	GGTGACGGCC	AGGACCACTG	AGGGCTTCGG	GCCAGAGCGC	5100
	GAGGGGATCA	TCACCATAGA	GTCCAGGAT	GGAGGACCTT	TCCCGCAGCT	GGGCGAGCGT	5160
	CGCGGGCTCT	TCCAGCACCC	GCTGCAAAAG	GAGTACAGCA	GCATCAACCAC	CACCCACACC	5220
	AGGCGGAGCG	AGCCCTCTCT	AGTGATGGGG	CCGACCTTGG	GGGCCAGCA	CCTGGAGGCA	5280
	GGCGGCTCCC	TCACCCGCGA	TGTGACCCAG	GAGTTTGTGA	GCCGGACACT	GACCAACAGC	5340
40	GGAAACCTTA	GCACCCACAT	GGACCAACAG	TTCTTCCAAA	CTTGACCGCA	CCCTGCCCCA	5400
	CCCCCGCCAT	GTCCCACTAG	CGGTCTCTCC	GACTCCTCTC	CCGAGAGCCT	CTCAGCTACT	5460
	CCATCCTTGC	ACCCCTGGGG	GCCAGGCCCA	CCCGCATGCA	CAGAGCAGGG	GCTAGGTGTC	5520
	TCCTGGGAGC	CATGAAGGGG	GCAAGGTCCG	TCCTCTGTGG	GCCCAAACTT	ATTGTAAACC	5580
45	AAAGAGCTGG	GAGCAGCACA	AGGACCCAGC	CTTTGTTCTG	CACCTAATAA	ATGTTTTTGC	5640
	TACTG						

Seq ID NO: 561 Protein sequence
Protein Accession #: NP_000204.1

50	1	11	21	31	41	51	
	MAGPRPSPWA	RLLLAALISV	SLSGTLANRC	KKAPVKSCTE	CVRVDKDCAY	CTDEMFRDRR	60
	CNTQAEALLA	GOQRESIVVM	BSSPQITEET	QIDTTLRRSQ	MSPGQLRVL	RPGEERHFH	120
	EVFEPLSEPV	DLYLMDFSN	SMSDDLNLK	KMGQNLARVL	SQLTSDYITG	PGKFVDKVS	180
	PQTDMPERKL	KEPWPNSDP	PSPKNVISLT	EDVDEFKNKL	QGERISGNLD	APGEGFDAIL	240
55	QTAVCTRDIG	WRPDSTHLLV	PSTESAFHYE	ADGANVLGAI	MSRNDERCHL	DTTGTYYQYR	300
	TQDYPVSPTL	VRLAKHNII	PIFAVTNYSY	SYEKLHTYF	PVSSLGVLE	DSSNIVELLE	360
	EAFNRIRSNL	DIRALDSPRG	LRTEVTSKMP	QKTRTGSFHI	RRGEVGIYOV	QLRALEHVDG	420
	THVCQLPEQD	KGNHILKPSF	SDGLKMDAGI	ICDVCTCELO	KEVRSARCSF	NGDFVCGQCV	480
60	CSBGWSGQCT	NCSTGSLSDI	QPCLREGEDK	PCSGRGECQC	GHCVCYGEGR	YEQGFCEYDN	540
	PQCPRTSGFL	CNDRGRCSMG	QCVCPEGWTG	PSCDCPLSNA	TCIDSNGGIC	NGRGHCBCGR	600
	CHCQQSLYT	DTICEINYSY	IHPGLCEDLR	SCVQCOAWGT	GEKKGRTECE	CNFKVKMVE	660
	LKRAEEVVVR	CSFRDEDDDC	TYSYTMEDGD	APGPNSTVLV	HKKKDCPPGS	FWWLIPLLLL	720
	LLPLLALLLL	LCWKYCACCK	ACLALLPCCN	RGHMVGFKED	HYMLRENLM	SDHLDTPMLR	780
65	SGNLKGRDVV	RKWTNNMQR	PGFATHAASI	NPTLVPIYGL	SLRLARLCTE	NLLKPDTRC	840
	AQLRQVEVEN	LNEVYRQISG	VHKLQQTQFR	QPNAGKKQD	HTIVDTVLMA	PRSAKPALLK	900
	LTEKQVEQRA	PHDLKVAPGY	YTLTADQDAR	GMVEFQEGVE	LVDVRVPLFI	RPEDDEKQL	960
	LVEAIDVPAG	TATLGRRLVN	ITIIKEQARD	VVSFEQPEFS	VSRGDQVARI	PVIRRVLDGG	1020
	KSQVSYRTQD	GTAGGNRDYI	PVEGELLFQP	GEAWKELQVK	LLELQEVDSL	LRGRQVRRFH	1080
70	VQLSNPKFGA	HLGQPHSTTI	IIRDFDELDR	SFTSQMLSSQ	PPPHGDLGAP	QNPNAKAAGS	1140
	RKIHFNWLPF	SGKPMGYRVK	YWIQDSESE	AHLDSKVPS	VELTNLYPYC	DYEMKVCAYG	1200
	AQGEQPYSSL	VSCRTHQVEP	SEPGRLAFNV	VSSTVTQLSW	AEPATNNGEI	TAYEVCYGLV	1260
	NDDNRPIGPM	KKVLVDNPKN	RMLLIENLRE	SQPYRYTVKA	RNGAGWGPER	EAIINLATQP	1320
	KRPMIPIIP	DIPIVDAQSG	EDYDSFLMYS	DDVLRSPSGS	QRPVSDDTE	HLVNGRMDFA	1380
75	FPGSTNSLHR	MTTTSAAAYG	THLSFHVPHR	VLSTSTLTTR	DYNSTLRSEH	SESTTLPRDY	1440
	STLTSVSSHD	SRLTAGVPDT	PTRLVFSALG	PTSLRVSWQE	PRCERPLQGY	SVEYQLLNGG	1500
	ELHRLINIPN	AQTSVVVEDL	LENHSYVFRV	RAQSQEGWGR	EREGVITIES	QVHPQSPICP	1560
	LPGSAFTLST	PSAPGLVFT	ALSPDSLQLS	WERPRRPNGD	IVGYLVTCM	AQGGGPATAP	1620
	RVDGSPESR	LTVPLGSENV	PKFRVQART	TEGFGPEREG	IITIESQDGG	PFPQLGSRAG	1680
80	LPQHPLQSEY	SSIITHTTSA	TEPFLVDGPT	LGAQHLEAGG	SLTRHVTQEF	VSRTLTTSGT	1740
	LSTMDQQPP	QT					

Seq ID NO: 562 DNA sequence
Nucleic Acid Accession #: NM_013332.1
Coding sequence: 1..63

85	1	11	21	31	41	51

GCACGAGGGC GCTTTTGTCT CCGGTGAGTT TTGTGGCGGG AAGCTTCTGC GCTGGTGCTT 60
AGTAACCGAC TTTCCTCCGG ACTCCTGCAC GACCTGCTCC TACAGCCGGC GATCCACTCC 120
CGGCTGTTC CCGGAGGGT CCAGAGGCCCT TTCAGAAAGG GAAGGCAGCT CTGTTTCTCT 180
GCAGAGGAGT AGGGTCCTTT CAGCCATGAA GCATGTGTGG AACCTCTACC TGTTAGGTGT 240
GGTACTGACC CTACTCTCCA TCCTCGTTAG AGTGATGGAG TCCCTAGAAG GCTTACTAGA 300
GAGCCCATCG CCTGGGACCT CCTGGACCAC CAGAAGCCAA CTAGCCAACA CAGAGCCAC 360
CAAGGCCCTT CCAGACCATC CATCCAGAAG CATGTGATAA GACCTCCTTC CATACTGGCC 420
ATATTTTGA AACTGACCT AGACATGTCC AGATGGGAGT CCCATTCTTA GCAGACAAGC 480
TGAGCACCGT TGTAAACGAG GAACTATTAC TAGGCCTTGA AGAACCTGTC TAACTGGATG 540
CTCATTGCCT GGGCAAGGCC TGTTTAGGCC GGTTCGGGTG GCTCATGCCT GTAATCCTAG 600
CACTTTGGGA GGCTGAGGTG GGTGGATCAC CTGAGGTGAG GAGTTCGAGA CCAGCCTCGC 660
CAACATGGCG AAACCCCATC TCTACTAAAA ATACAAAAGT TAGCTGGGTG TGGTGGCAGA 720
GGCCTGTAAT CCCAGTTCCT TGGGAGGCTG AGGCGGGAGA ATTGCTTGAA CCGCGGGACG 780
GAGGTTCGAG TGTAAACGAG TCGCACTGCT GTACCCAGCC TGGGCCACAG TGCAAGACTC 840
CATCTCAAAA AAAAAAAGAA AAAAAAAGC CTGTTTAATG CACAGGTGTG AGTGGATTGC 900
TTATGGCTAT GAGATAGGTT GATCTCGCCC TTACCCGGGG GTCTGGGTGA TGCTGTGCTT 960
TCCTCAGCAG TATGGCTCTG ACATCTCTTA GATGTCCCAA CTTGAGCTGT TGGGAGATGG 1020
TGATATTTTC AACCTACTT CCTAACATC TGTCTGGGGT TCCTTTAGTC TTGAATGTCT 1080
TATGCTCAAT TATTTGGTGT TGAGCCTCTC TTCCACAAGA GCTCCTCCAT GTTTGGATAG 1140
CAGTTGAAGA GGTGTGTGGG GTGGGCTGTT GGGAGTGAGG ATGGAGTGTG CAGTGCCCAT 1200
TTCTCATTTT ACATTTTAAA GTCTGTCCTC CAACATAGTG TGTATTGGTC TGAAGGGGGT 1260
GGTGGGATCG CAAAGCCTGC TCAAGTTATG GACATTGTGG CCACCATGTG GCTTAAATGA 1320
TTTTTCTAA CTAATAAAGT GGAATATATA TTTCAAAAAA AAAAAA AA

25 Seq ID NO: 563 Protein sequence
Protein Accession #: NP_037464.1

1 11 21 31 41 51
MKHVLNLYLL GVLVTLISIP VRVMESLEGL LESPSPGTSH TTRSQLANTE PTKGLPDRPS 60
RSM

35 Seq ID NO: 564 DNA sequence
Nucleic Acid Accession #: NM_023915.1
Coding sequence: 250..1326

1 11 21 31 41 51
GGCACGAGGG TTTCGTTTTC ATGCTTTACC AGAAAAATCCA CTTCCCTGCC GACCTTAGTT 60
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120
GTGAATGGAC AGGCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180
CCACGCTCTC AATCGTCCCC AAGTGTTCCT TGACACGCAT CTTGCTTAC AGTGCAATC 240
AACTGAAGAA TGGGGTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCACGGC 300
CAAGAGAGTG ACATTCAGG CAACAGGAGC GACGGGGCAG GAAAGAACAC CACCTTTCAC 360
AATGAATTG ACACAATTGT CTTGCGGTG CTTTATCTCA TTATATTGT GGCAGCATC 420
TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAC CAGCTTCATA 480
TTCTATCTCA AAAACATAGT GGTTCAGAC CTCATAATGA CGTGACATT TCCATTTCGA 540
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600
TCAGTTTGT TTATGCAAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCAIT 660
GATCGCTATC TGAAGGTGGT CAAGCCATT GGGGACTCTC GGATGTACAG CATAACCTTC 720
ACGAAGGTTT TATCTGTTTG TGTTTGGGTG ATCATGGCTG TTTTGTCTTT GCCAACATC 780
ATCTCAGCAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAAA ACTTAAAGT 840
CCTTTGGGGG TCAAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900
GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960
AGGCAATTCA TAAGTCAGT AAGCGGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT 1020
GTGGCTGTGT TTTTACCTG CTTTCTACCA TATCACTGT GCAGAAATCC TTTTACTTTT 1080
AGTCACTTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCTATATTA CTGCAAGAA 1140
ATTACACTTT TCTGTCTGCG GTGTAATGTT TGCCTGGATC CAATAATTTA CTTTTCATG 1200
TGTAAGTCAT TTCAAGAAG GCTGTTCAAA AAATCAAATA TCAGAACCCAG GTAGGAAAGC 1260
ATCAGATCAC TGCAAGAGTG GAGAAGATCG GAAGTTGCGA TATATTATGA TTACTGAT 1320
GTGTAGGCTT TTTATGTTT GTTGAATCG ATATGTACAA AGTGTAATAA AATGTTTCTT 1380
TTCATTATCC TTAATAAAAA AA

65 Seq ID NO: 565 Protein sequence
Protein Accession #: NP_076404

1 11 21 31 41 51
MGPNLTAKL PNNELHQES HNSGNRSDGP GQNTLHNEF DTIVLPVLYL IIFVASILLN 60
GLAVWIFFHI RNKTSFIFYL KNIVVADLIM TLTPFFRIVE DAGFGPWYFK FILCRYTSVL 120
PYANMYSIV FLGLISIDRY LKVVKPFQDS RMYSTFTKV LSVCVWVIMA VLSLPNIILT 180
NGQPTEDNIH DCSKLKSPFG VKWHTAVTVV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQP 240
ISQSSRRKHE NQSIKRVVAV FFTCFLPYHL CRIPPTPSHL DRLLDESAQK ILYYCKBITL 300
FLSACNVCLD PIYFFMCRS PSRLPKKSN ITRSESIRS LQSVRRSEVR IYYDYTDV

75 Seq ID NO: 566 DNA sequence
Nucleic Acid Accession #: NM_005365.1
Coding sequence: 1..948

80 1 11 21 31 41 51
ATGTCTCTCG AGCAGAGGAG TCCGCACTGC AAGCCTGATG AAGACCTTGA AGCCCAAGGA 60
GAGGACTTGG GCTGTATGGG TGCACAGGAA CCCACAGGCG AGGAGGAGGA GACTACCTCC 120
TCCTCTGACA GCAAGAGAGA GGAGGTGTCT GCTGCTGGGT CATCAAGTCC TCCCAGAGT 180
85 CCTCAGGAG GCGCTTCTCT CTCCATTTC GTCTACTACA CTTTATGGAG CCAATTGAT 240
GAGGCTCCA GCACTCAAGA AGAGGAAGAG CCAAGCTCCT CGGTGACCC AGCTCAGCTG 300
GAGTTCATGT TCCAAGAAGC ACTGAAATG AAGGTGGCTG AGTTGGTTCA TTTCTGCTC 360

5
10
CACAAATATC GAGTCAAGGA GCCGGTCACA AAGGCAGAAA TGCTGGAGAG CGTCATCAAA 420
AATTACAAGC GCTACTTTCC TGTGATCTTC GGCAAAGCCT CCGAGTTTCAT GCAGGTGATC 480
TTTGGCACTG ATGTGAAGGA GGTGGACCCC GCGCGCCACT CCTACATCCT TGTCACTGCT 540
CTTGGCCCTC OGTGCGATAG CATGCTGGGT GATGGTCATA GCATGCCCAA GGCGGCCCTC 600
CTGATCATTG TCCTGGGTGT GATCCTAACC AAAGACAAC TCGCCCTCGA AGAGGTTATC 660
TGGGAAGCGT TGAAGTGTAT GGGGGTGTAT GTTGGGAAG AGCACATGTT CTACGGGGAG 720
CCGAGGAAGC TGCTACCCCA AGATTGGGTG CAGGAAAAC TACCTGGAGTA CCGGCAGGTG 780
CCGCGCAGTG ATCTGCGCA CTACGAGTTC CTGTGGGGT CCAAGGCCCA CGCTGAAACC 840
AGCTATGAGA AGGTCAATAA TTATTTGGTC ATGCTCAATG CAAGAGAGCC CATCTGCTAC 900
CCATCCCTTT ATGAAGAGGT TTTGGGAGAG GAGCAAGAGG GAGTCTGA

Seq ID NO: 567 Protein sequence
Protein Accession #: NP_005356.1

15
20
1 11 21 31 41 51
MSLEQRSPHC KPDELEAQQ EDLGLMGAQE PTGEEEEETS SSSDSKEEVS AAGSSSPFQS 60
PQAGASSSIS VYVTLWSQFD EGSSSQEEEE PSSSVDPALQ EFMFQEAALK KVALVHFL 120
HKYRVKEPVT KAEMLESVIK NYKRYPPVIP GKASEFMQVI PGTDVKEVDP AGHSYILVTA 180
LGLSCDSMLG DGHSMFKAAL LIIVLGVILT KDNCAPEEVI WEALSVMGVY VGKEHMFYGE 240
PRKLLTQDWF QENYLEYRQV PGSDPAHYEF LWGSKAHAE TSYEKVINYL VMLNAREPICY 300
PSLYEEVLGE EQEYV

25
Seq ID NO: 568 DNA sequence
Nucleic Acid Accession #: NM_014400
Coding sequence: 86..1126

30
35
40
45
50
55
60
1 11 21 31 41 51
GGTACTCAT CCTGGGCTCA GGTAAGAGGG CCCGAGCTCG GAGGCGGCAC ACCCAGGGGG 60
GAGCCCAAGG GAGCAGGACG GAGCCATGGA CCCCGCCAGG AAAGCAGGTG CCCAGGCCAT 120
GATCTGGAGT GCAGGCTGGC TGCTGCTGCT GTGCTTCGC GGAGGAGCGC AGGCCCTGGA 180
GTGCTACAGC TGCCTGCAGA AAGCAGATGA CGGATGTCC CCGAACAGA TGAAGACAGT 240
GAAGTGCGCG CCGCGCGTGG ACCTGTGCAC CGAGGCGGTG GGGGCGGTGG AGACCATCCA 300
CGGACAAATC TCCTGTGCAG TGCSGGGTG CGGTTCGGGA CTCGCCGGCA AGAATGACCG 360
CGGCTTGAAT CTTACCGGGC TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420
CTGCAACGCC AAGCTCAACC TCACCTCGCG GCGCTCGAC CCGGCAGGTA ATGAGAGTGC 480
ATACCCGCCC AACCGCGTGG AGTGCTACAG CTGTGTGGGC CTGAGCCGGG AGGCGTGCCA 540
GGGTACATCG CCGCGGCTCG TGAGCTGCTA CAACGCCAGC GATCATGTCT ACAAGGGCTG 600
CTTCGACGGC AACGTCACTT TGACGGCAGC TAATGTGACT GTGTCTTGG CTGTCCGGGG 660
CTGTGTCCAG GATGAATCTT GCACTCGGGA TGGAGTAACA GGCCCGAGGT TCACGCTCAG 720
TGCTCTCTGT TGCCAGGGGT CCGCTGTAA CTCTGACCT CCGAACAGA CTTACTTCTC 780
CCCTCGAATC CCACCCCTTG TCCGGCTGCC CCCTCCAGAG CCCACGACTG TGGCTCAAC 840
CACATCTGTC ACCACTTCTA CCTCGGCCCC AGTGAGACCC ACATCCACCA CCAAAACCAT 900
GCCAGGCCCA ACCAGTCAGA CTCGAGACA GGGAGTAGAA CACGAGGCTT CCGGGATGA 960
GGAGCCGAGG TTGACTGGAG GCGCGCTGG CCACCGAGAC CGCAGCAATT CAGGGCAGTA 1020
TCCTGCAAAA GGGGGGCCCC AGCAGCCCCA TAATAAGGC TGTGTGGCTC CCACAGCTGG 1080
ATTGGCAGCC CTTCCTTTGG CCGTGGCTGC TGGTGTCTTA CTGTGAGCTT CTCACCTGG 1140
AAATTTCCCT CTCACCTACT TCTCTGGCCC TGGGTACCCC TCTTCTCATC ACTTCTGTT 1200
CCCAACCATG GACTGGGCTG GCCCAGCCCC TGTTTTCCA ACATTCGCCA GTATCCCCAG 1260
CTTCTGCTGC GCTGTGTTGC GGCITTTGGA AATAAAATAC GSTTGTATAT ATTCTGGCAG 1320
GGGTGTTCTA GCTTTTGGAG GACAGCTCCT GTATCCTTCT CATCCTTGTC TCTCCGCTTG 1380
TCCTCTTTTG ATGTAGAGAC AGAGTGAGAG AAGTCAGCTG TCACGGGGAA GGTGAGAGAG 1440
AGGATGCTAA GCTCTCACT CACTTCTCC TAGCCAGCCT GGACTTTGGA GCGTGGGGTG 1500
GGTGGGACAA TGGCTCCCCA CTCTAAGCAC TGCCCTCCCT ACTCCCCCA TCTTTGGGGA 1560
ATCGGTTCCC CATATGTCTT CCTTACTAGA CTGTGAGCTC CTCGAGGGCA GGGACCGTGC 1620
CTTATGTCTG TGTGTGATCA GTTCTGGCCA CATAAATGCC TCAATAAGA TTAAATTACT 1680
TTGTATAGTG AAAAAAAA

60
Seq ID NO: 569 Protein sequence
Protein Accession #: NP_055215

65
70
1 11 21 31 41 51
MDPARKAGAQ AMIWAGWLL LLLLRRGAQA LECYSCVOKA DDGCSNPKMK TVKCAPGVDV 60
CTEAVGAVET IHGFSLAVX GCGSLPGKN DRGLDLHGLL AFILQQAQCA DRNNAKLNLT 120
SRALDPAGNE SAYPPNGVEC YSCVGLSREA CQGTSPPVVS CYNASDHVYK GCFDGNVTLT 180
AANVTYSLPV RGCVDPEFCT RDGVTGPFT LSGSCCQGSR CNSDLRNKTY FSPRIPLVR 240
LPPPEPTTVA STTSVTSTTS APVRPTSTTK PMPAPTSQTP RQGVHEASR DEEPRLTGGA 300
AGHQDRSNSG QYPAKGGPQ PHNKGCVAPT AGLAALLLAV AAGVLL

75
Seq ID NO: 570 DNA sequence
Nucleic Acid Accession #: NM_005329.1
Coding sequence: 1..1662

80
85
1 11 21 31 41 51
ATGCCGGTGC AGCTGACGAC AGCCCTGCGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60
GTGCTGGGTG GCATCTCTGC AGCCTATGTG ACGGCTACC AGTTCATCCA CACGGAAAAG 120
CACTACTCTG CCTTCGSCCT GTACGGCGCC ATCCTGGGCC TGCACCTGCT CATTAGAGC 180
CTTTTGGCCT TCTTGGAGCA CCGCGCATG CGACGTGCCG GCCAGGCCTT GAAGCTGCCC 240
TCCCCCGCGC GGGGCTCGGT GGCACGTGTC ATTGCGCGT ACCAGGAGGA CCGTGACTAC 300
TTGGCGCAAGT GCCTGCGCTC GGCCCGAGCG ATCTCCTTCC CTGACCTCAA GGTGTCATG 360
GTGGTGGATG GCAACGCCCA GGAGGACGCC TACATGCTGG ACATCTTCCA CGAGGTGCTG 420
GGCGGCACCG AGCAGGCCG CTCTTTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG 480
GGTGAGACGG AGGCCAGCCT GCAGGAGGGC ATGGACCGTG TGCGGGATGT GGTGCGGGCC 540
AGCACCTTCT CGTGATCAT GCAGAAAGTG GGAGGCAAGC GCGAGGTCAT GTACACGGCC 600

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

TTCAAGGCC TCAGCGATTC GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660
 GATCCAGCCT GCACCATCGA GATGCTTCGA GTCTGGAGG AGGATCCCCA AGTAGGGGGA 720
 GTGCGGGGAG ATGTCCAGAT CCTCAACAAG TACGACTCAT GGATTTCCCT CCTGAGCAGC 780
 GTGCGGTACT GGATGGCCTT CAACGTGGAG CGGGCCTGCC AGTCTACTT TGGCTGTGTG 840
 CAGTGTATTA GTGGGCCCTT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCTCGGAG 900
 GACTGTATCC ATCAGAAGTT CCTAGGCAGC AAGTGCAGCT TCGGGGATGA CCGGCACCTC 960
 ACCAACCGAG TCCTGAGCCT TGGCTACCGA ACTAAGTATA CGCGCGCTC CAAGTGCCTC 1020
 ACAGAGACCC CCACTAAGTA CCTCCGGTGG CTCAACAGC AAACCCGCTG GAGCAAGTCT 1080
 TACTTCCGGG AGTGGCTCTA CAACTCTCTG TGGTTCCATA AGCACCCT CTGGATGACC 1140
 TAGAGTCAG TGGTCACGGG TTTCTTCCCC TTCTTCTCTA TTGCCACGGT TATACAGCTT 1200
 TTCTACCGGG GCGCATCTG GAACATTCTC CTCTTCTCTG TGACGGTGCA GCTGGTGGGC 1260
 ATTATCAAGG CCACCTACGC CTGCTTCCCT CGGGGCAATG CAGAGATGAT CTTATGTTC 1320
 CTCTACTCCC TCCTTATATG GTCCAGCCTT CTGCGCGCCA AGATCTTTGC CATTGTCTAC 1380
 ATCAACAAAT CTGGCTGGGG CACCTCTGGC CGAAAAACCA TTGTGTGAA CTTATTGGC 1440
 CTCATTCTCT TGTCCATCTG GGTGGCAGTT CTCCTGGGAG GGTGGCCTA CACAGCTTAT 1500
 TGCCAGGACC TGTTCAGTGA GACAGAGCTA GCCTTCTCTG TCTCTGGGGC TATCTGTAT 1560
 GGCTGTCTCT GGTGGCCTT CCTCATGCTA TATCTGGCCA TCATCGCCCG GCGATGTGGG 1620
 AAGAAGCCGG AGCAGTACAG CTGGCTTTT GCTGAGGTGT GA

Seq ID NO: 571 Protein sequence
 Protein Accession #: NP_005320.1

1 11 21 31 41 51
 MFVQLTTLALR VVGTSFLFALA VLGGILAAYV TGYQFIHTEK HYLSFGLYGA ILGLHLLIQS 60
 LFAFLQEHRRM RRAGQALKLP SPRRGSVLALC IAAVQEDPDY LRKCLRSAQR ISFPDLKVVV 120
 VVDGNRQEDA YMLDIFHEVL GGTGQAGFFV WRSNPFHEAGE GETEASLQEG MDRVRDVVRA 180
 STFSCTIMQKW GSKREVMYTA FKALGDSVDY IQVCDSDTVL DPACTIEMLR VLEEDPQVGG 240
 VGGDVQILNK YDSWISFLSS VRYWMAFNV E RACQSYFGCV QCISGPLGMY RNSLLQGFLE 300
 DWYHQKPLGS KCSFGDDRLH TNRVLSLGR TKYTARSKCL TETPTKYLRL LNQQTRWSKS 360
 YFREMLYNLS WFKFHEHLMNT YESVVTGFFP FFLIATVIQL FYRGRIWNIL LFLLTQVLVG 420
 IIKATYACFL RGNAEIIFMS LYSLLYMSLS LPAKIFAIAI INKSGWGTSG RKTIVNVNFI 480
 LIPVSIWVAV LLGLLAYTAY QDLFSETEL AFLVSGAILY GCYVWALLML YLAIIARRCG 540
 KKPEQYSLAF AEV

Seq ID NO: 572 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 148-7095

40
 45
 50
 55
 60
 65
 70
 75
 80
 85

1 11 21 31 41 51
 CACACATACG CACGCACGAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAAAC ATTTCTCTCG CTCCCTCTCC CTCTCCATC TGAGAAGCAG AGGAGCCGCA 120
 CGCGCAGGGG CCGCAGACCG TCTGGAAATG CGAATCTTAA AGCGTTTCCT CGCTTGCAAT 180
 CAGCTCCTCT GTGTTTGGCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTGTGTTGAAG AGATTTGGCTG GTCTATACCA GGAGCAGCTGA ATCAAAAAAA TTGGGGAAG 300
 AATATCCCAA CATGTAATAG CCCAAACAAA TCTCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAACTTAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA 420
 AACACATTTA TTCAATACAC TGGGAAAAACA GTGGAAATTA ATCTCACTAA TGACTACCGT 480
 GTCAGCGGAG GAGTTTTCGA AATGGTGTIT AAAGCAAGCA AGATAACTTT TCACCTGGGA 540
 AATGCAATA TGTCTATCTG TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600
 GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660
 GGAAGAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAAATTG 720
 GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTG GTTTTGAGGA GCAGGCTGCT 780
 TTAGATCCAT TCATATCTGT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTTAC 840
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAG CCACTTGGCT GTTTTGTGT AAGTTCTTAC AATGCAACAA 960
 TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTGAGAGCA ACAGTACAAG 1020
 TTCTCTAGAC AGGTGTTTTC CTCAATACAT GGAAGAGGAG AGATTCTATGA AGCAGTTTGT 1080
 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAAAT ATACCAGCCT TCTTGTTACA 1140
 TGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGAGT TTTGTACCAG 1200
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATTG TCAATAATTG GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTCGACAT GCCTACTGAT 1380
 AATCCTGAAC TTGATCTTTT CCTGGAATTA ATTGGAACAG AAGAAATAAT CAAGGAGGAG 1440
 GAAGAGGGAA AAGACATGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500
 AACCAATCA GGAAGAAAGG ACCCCAGATT TCTACCAAA CACACTACAA TCGCATAGGG 1560
 ACGAATACCA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTTGGA 1620
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
 ACAGAAAAAG ATATTTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACAG TTCTTAGATC TCCACATATG 1800
 AACTTGTGCG GGAATGACGA ATCTTTAAAT ACAGTTTCTA TAACAGAAAT TGAGGAGGAG 1860
 AGTTTATTGA CCAATTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920
 GCAACTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980
 GAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040
 GAAGATTCAA CTTCATCAGG TTCAGAAAG TCACTAAAGG ATCCTTCTAT GGAGGGAAT 2100
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160
 AGCTTTCTCC AGACTTAATTA CACTGAGATA CGTGTGTATG AATCTGAGAA GACAAACAG 2220
 TCCTTTTCTG CAGGCCAGCT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280
 CATTATTCTA CTTTGCCTA TTTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340
 TCCAGACAA AGGATTTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG 2400
 GTATACAATG GTGAGACACC TCTTCAACCT TCCTACAGTA GTGAAGTCTT TCCTCTAGTC 2460
 ACCCTTTTGT TGCTTGACAA TCAGATCCTC AACACTACCC CTGCTGCTTC AAGTAGTGAT 2520
 TCGSCCTTGC ATGCTACGCC TGTATTTCCT AGTGTGATG TGTCAATTGA ATCCATCCTG 2580
 TCTTCTATG ATGGTGACCC TTTGCTTCCA TTTCTCTG CTTCCTTCTG TAGTGAATTG 2640
 TTTCCGCATC TGCATACAGT TTCTCAAATC CTTCCACAG TTAATTCTAGC TACCGAGAGT 2700
 GATAAGGTGC CCTTGATGCT TTCTTGCCA GTGGCTGGG GTGATTGCT ATTAGAGCCC 2760

	AGCCTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGGTAGTG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACCACCC	2880
	AGCAGTGATG	CCATGATGCA	TGCACGTTCT	TCAGGGCCCTG	AACTTCTCTA	TGCTTGTGCT	2940
5	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTTTCTTACA	GTTCTGCAAT	ACCTGTGSCAT	3000
	GATTCTGTGG	GTGTAACCTA	TCAGGTTTCC	TTATTTAGCG	GCCTTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AATGTCCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCACTT	GTTTCTGTAG	CTGAATTTAC	ATATACACAA	3240
10	TCGTGTTTGG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGGAATGAG	3300
	ACTGAACCTG	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGGCAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAT	AACTTTTCAG	TTCAACCTAC	ACATACCTGT	3540
	TCTCAAGCAT	CTGGTGACAC	TTGCTTAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
15	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAT	TGAAGTATTG	CTACAACCTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAACATGT	TCTTCCAGCT	GTGCCAGTG	ATCCAAATAT	GGTTGAAACC	3780
	CCCAAGTGTG	ATAAAATCTG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
20	AGTGAAACA	TGCTGCACTC	TACATCTGTA	CCAGTTTITG	ATGTGTCCGC	TACTTCTCAT	3900
	ATGCACTCTG	CTTCACTTCA	AGGTTTGACC	ATTTCTCTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTTTGTGTAA	AAAGTGAAAG	TTCCCAACCA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAA	CAGGCCCATC	CCCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AAATTGATGAA	CCATTAAATA	CACATAAA	TAAGCTTATA	4140
25	CATTCCGATG	AAATTTTAA	CTCCACCAAA	AGTTCTGTGA	CTGTAAAGT	ATTTGCTGGT	4200
	ATTCCAACAG	TGCTTCTGA	TACATTGTGA	TCTACTGATC	ATTTCTGTTCC	TATAGGAAAT	4260
	GGGCATGTTG	CACTTACAGC	TGTTTCTCCC	CACAGAGATG	GTCTGTAAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
30	AGAGGTAGTG	ATGGCTTATC	CATTCTAAG	TGTATGTCAT	GCTCATCCTA	TAGAGAATCA	4500
	CAGGAAAGG	TAATGAATGA	TTCAAGACCC	CAGGAAACA	GTCTTATGGA	TCAGAATAAT	4560
	CCAATCTCAT	ACTCACTATC	TGAGAAATCT	GAAGAAGATA	ATAGAGTCAC	AAGTGTATCC	4620
	TCAGACAGTG	AAACTGGTAT	GGACAGAAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGGTA	4680
	TOCCAAAAGC	ACAATGATGG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
35	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCTGACAA	GTGATGAAGA	AAGTGTGATCA	4800
	GGGCAAGTTC	CCTCAAGTAG	CCTTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTGTGA	4860
	GACACTAATG	AAAAGATGTC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AATAACTCTCT	4920
	GGATTCCAC	AGTCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAAGT	GTTCCACGTT	4980
	TCAGAGGCG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTG	5040
40	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	5100
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCACACC	TATCTTTCCA	5220
	ATTTAGATG	ATGTGGAGC	AATCCAATA	AAGCACTTTC	CAAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	5340
45	CAGAGCTGTA	CTGTGTACTT	AGGTATTACA	GCAGACAGCT	CCAACCAACC	AGACAACAAG	5400
	CACAAGATGA	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	5520
	AACAGACCAC	AGCTTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTC	5580
	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	5640
50	AAAGGAAGGA	AAAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	5700
	TTTCTGGTCA	CTCAGAAAGT	TGTGCAAGTG	CTTGCCCTATT	ATACTGTGAG	GAATTTTACT	5760
	CTAAGAAACA	AAAAATATAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	5820
	ACACAGTATG	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	5880
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTGCTC	5940
55	CACCTGAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGACG	6000
	CAGATTCAAC	AGCAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAACAACAT	CCGTTCACAA	6060
	AGAAATTTAT	TGATCAAAAC	TGAGGAGCAA	TATGTCTTCA	TTCAATGATC	ACTGGTTGAG	6120
	GGCACTATTA	GTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCTAGCCTA	TGTTAATGCA	6180
	CTCCTCATTC	CTGGACCAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GCTCCTGAGC	6240
60	CAGTCAAAAT	TACAGCAGAG	TGACTATTCT	GCAGCCCTAA	AGCAATGCAA	CAGGGAAGAG	6300
	AATCGAATCT	CTTCTATCAT	CCCTGTGGAA	AGATCAAGGG	TGGCATTTTC	ATCCCTGAGT	6360
	GGAGAAAGCA	CAGACTACAT	CAATGCCTCC	TATATCATGG	GCTATTACCA	GAGCAATGAA	6420
	TTTATCATTA	CCCAGCACCC	TCTCCTTCAT	ACCATCAAGG	ATTCTGGAG	GATGATATGG	6480
	GACCATTAATG	CCCAACTGGT	GGTTATGATT	CCTGATGGCC	AAAACATGGC	AGAAGATGAA	6540
65	TTTGTTTACT	GGCCAAATAA	AGATGAGCCT	ATAAATTTGT	AGAGCTTTAA	GGTCACCTTT	6600
	ATGGCTGAAG	AACACAAATG	TCTATCTAAT	GAGGAAAAAC	TTATAATTCA	GGACTTTATC	6660
	TTAGAGGCTA	CACAGGATGA	TTATGTACTT	GAGTGAGGC	ACTTTCAGTG	TCCTAAATGG	6720
	CCAAATCCAG	ATAGCCCCAT	TAGTAAACT	TTTGAACCTA	TAAGTGTATT	AAAAGAAGAA	6780
	GCTGCCAATA	GGGATGGGCC	TATGATTGTT	CATGATGAGC	ATGGAGGAGT	GACGGCAGGA	6840
70	ACTTTCTGTG	CTCTGACAAC	CCTTATGCAC	CAACTAGAAA	AAGAAAATTC	CGTGGATGTT	6900
	TACCAAGTAG	CCAAGATGAT	CAATCTGATG	AGGCCAGGAG	TCTTTGCTGA	CATTGAGCAG	6960
	TATCAGTTTC	TCTACAAAGT	GATCCTCAGC	CTTGTGAGCA	CAAGGCAGGA	AGAGAATCCA	7020
	TCCACCTCTC	TGGACAGTAA	TGGTGACGCA	TTGCCGTGAT	GAATATAGC	TGAGAGCTTA	7080
	GAGTCTTTAG	TTTAAACACG	AAAGGGGTGG	GGGGACTCAC	ATCTGAGCAT	TGTTTTCTCT	7140
75	TTCTTAAAT	TAGGCAGGAA	AATCAGTCTA	GTTCTGTTAT	CTGTTGATT	CCCATCACCT	7200
	GACAGTAAT	TTTATGACAT	AGGATTCTGC	CGCCAAATTT	ATATCATTTA	CAATGTGTGC	7260
	CTTTTGTCAA	GACTTGTAAT	TTACTTATTA	TGTTTGAAC	AAAATGATTG	AATTTTACAG	7320
	TATTTCTAAG	AATGGAATGT	TGGTATTTTT	TTCTGTATTG	ATTTTAAACG	AAAATTTCAA	7380
	TTTATAGAGG	TTAGGAATTC	CAAACTACAG	AAAAATGTTG	TTTTAGTGT	CAAAATTTTA	7440
80	CGTGATTGTT	TAGCAATAT	CAGGTTTGCT	AGAAATATAA	CTTTTAAATC	AGTAGCCCTG	7500
	AAATAAAACA	CTCTCCATA	TGATATTCAA	CATTTTACAA	CTGCAGTATT	CACCTAAAGT	7560
	AGAAATAATC	TGTTACTTAT	TGTAATATCT	GGCCTAGTGT	CTCCATGGAC	CAAAATTTATA	7620
	TTTATAATGT	TAGATTTTTA	TATTTTACTA	CTGAGTCAAG	TTTTCTAGTT	CTGTGTAATT	7680
	GTTTAGTTTA	ATGAGCTTAT	TCATTAGCTG	GTCTTACTCT	ACCAATTTTC	TGACATTGTA	7740
	TTGTGTTACC	TAAGTCAATTA	ACTTTGTTTC	AGCATGTAAT	TTTAACTTTT	GTGGAAAAATA	7800
85	GAAATACCTT	CATTTTGAAA	GAAGTTTTTA	TGAGAATAAC	ACCTTACCAA	ACATTGTTCA	7860
	AATGGTTTTT	ATCCAAGGAA	TTGCAAAAAT	AAATATAAAT	ATTGCCATTA	AAAAAAAATA	7920
	AAAAAAAATA	AAAAAAAATA	AAAA				

Seq ID NO: 573 Protein sequence:
Protein Accession #: Eos sequence

5	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQR	KLVEEIGWSY	TGALNQXNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNKKL	KFGQWDTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
	FKASRITFW	GKCNMSSDGS	EHSLEGQKFP	LEMQIYCPDA	DRFSSPEEAV	KGKGLRALS	180
10	ILFEVGTEN	LDPKALIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDMWIFK	DTVSISESQL	AVFCEVLTMQ	QSGYVLMMDY	LQNNFREQQY	KFSRQVFSY	300
	TGKEEIHFAV	CSSEPEVQA	DPENYTSLLV	TWERPRVVDY	TMIEKPAVLV	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNLLP	NMSYVLQIVA	ICTNGLYGKY	SDQLIVDMPT	DNPELDLPE	420
	LIGTEEIIKE	EEEGKIDIEG	AIVNPGRDSA	TNQIRKKEPO	ISTTTHYNRI	GTKYNEAKTN	480
15	RSPTRGSEPS	GKQDVPTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPIHT	VEGTASASLND	540
	GSKTVLRSPH	MNLSTGTAESL	NTVSIITEYE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPTIITYD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NVWFPSTDI	660
	TAQPPVSGGR	ESFLQNTYTE	IRVDESEKTT	KSPSAGPVMS	QGPSVTDLEM	PHYSTFAYFP	720
	TEVTPHAFTP	SSRQQLVST	VNVVYSQTTQ	PVYNGETPLQ	PSYSSEVPPL	VTPLLLDNQI	780
20	LNTTPAASSS	DALHATPVP	PSVDVSFESI	LSSYDGAPLL	PFSSASFSSE	LFRHLHTVSQ	840
	ILPQVTSATE	SKRVPLHASL	PVAGGDLLE	PSLAQYSDVL	STTHAASETL	EFGSESGVLY	900
	KLTFSPSQVEP	PSSDMMHAR	SSGPEPSYAL	SDNEGSQHIF	TVSYSSAIPV	HDSVGVITYQG	960
	SLFSGPSHIP	IPKSLITPT	ASLLQPTHAL	SGDGEWSGAS	SDSEFLLPDT	DGLTALNISS	1020
	PVSVAEFTYT	TSVFGDDNKA	LSKSEIYYGN	ETELQIPSPN	EMVVPSESTV	MPNMYDNVKN	1080
25	LNASLQETSV	SISSTKGMFP	GSLAHTTKV	FDHEISQVPE	NNFSVQPTH	VSQASGDTSL	1140
	KPVLSANSEP	ASSDPASSEM	LSPSTQLLFP	ETSASFSTEV	LLQPSFQASD	VDTLKLTVP	1200
	AVPSDPLIVE	TPKVKIKIST	MLHLIVNSNA	SSENMLHSTS	VPFVDSPTS	BMHSASLQGL	1260
	TISYASEKYE	FVLLKSESSH	QVVPFLYSND	ELFQTANLEI	NQAHPPKGRH	VFATFVLISD	1320
	EPLNTLINKL	IHSDEILTST	KSSVTGKVFA	GIPTVASDTF	VSTDHSPVIG	NGHVAITAVS	1380
30	PHRDGSVTSV	KLLFPFSKATS	ELSHSAKSDA	GLVGGGEDGD	TDDGDGDDDD	DRGSDGLSIH	1440
	KCMSCSSYRE	SQEKVMNDSD	THENSMDQON	NPISYSLSEN	SEEDNRVTSV	SSDSQTMGDR	1500
	SPGKSPSANG	LSQKNDGKE	ENDIQTGSAL	LPLSPESKAW	AVLTSDEESG	SGQGTSDSLN	1560
	ENETSTDFSF	ADTNEKDADG	ILAAGDSEIT	PGFPQSPTSS	VTSENSEVFH	VSEAEASNSS	1620
	HESRIGLAEG	LESEKKAIVP	LVIVSALTFI	CLVVLVGILI	YWRKCPQTAH	PYLEDTSTPR	1680
35	VISTPTPIFP	PISDVGAIPI	IKHFPKHVAD	LHASSGFTEE	PETLKEFYQE	VQSCVTDLGI	1740
	TADSSNHDPN	KHKNYRINIV	AYDHSRVKLA	QLAEKDGKLT	DYINANYVDG	YNRPKAYIAA	1800
	QGLKSTADE	FWRMTWEHNV	EIVIMITNLV	EKGRRKCDQY	WPADGSEEEY	NFLVTQKSVQ	1860
	VLAYYYTRNF	TLRNTKIKKG	SQKGRPSGRV	VTQYHYTQWP	DMGVPEYSLP	VLTFVRKAAV	1920
	AKRHAVGPV	VBCSAGVGR	GTIVLDSML	QQIQHSGTVN	IFGFLKHIRS	QRNVLVQTEE	1980
40	QVVFHDTLV	EAILSKETE	LDSEIHAYVN	ALLIPGPAGK	TKLEKQFQLL	SQSNIQOSDY	2040
	SAALKQCNRE	KNRTSSIIPV	ERSRVGISSL	SGEGTDYINA	SYIMGYQSN	EFIIITQHPLL	2100
	HTIKDFWRI	WDHNAQLVVM	IPDQNMAD	EPVYVPMKDE	PINCESFKVT	LMAEHKKCLS	2160
	NEEKLIQDF	ILEATQDDYV	LEVRHFQCPK	WPNPDSPISK	TFELISVIKE	EAAANROGPMI	2220
45	VHDEHGGVTA	GTFCALITLM	HQLEKENSVD	YVQVAKMINL	MRPGVPADIE	QYQFLYKVL	2280
	SLVSTRQEN	PSTSLDSNGA	ALPDGNIAES	LESILV			

Seq ID NO: 574 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 148-4518

50	1	11	21	31	41	51	
	CACACATACG	CACGACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTCTCTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
55	CGCGAGAGGG	CGCGAGAGCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCTC	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAA	AGATTGCGTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TGGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
60	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCAATGGAA	420
	AACACATTC	TTCATAACAC	TGGGAAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCTT	480
	GTCAGCGGAG	GAGTTTCAGA	AATGGTGTTT	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
65	GGAAAGAGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTT	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTAGTGC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATCTGTTT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTGTGTG	AAGTTCCTAC	AATGCAACAA	960
70	TCTGTTATG	TCACTGCTAT	GGACTACTTA	CAAAACAATT	TTCGAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAG	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTGTATCCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
75	GGTGCTATT	TCAATAATTT	GCTACCOAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCATTAATG	GCTTATATGG	AAAAACAGC	GACCAACTGA	TTGTGCACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACGT	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TGCGATAGGG	1560
80	ACGAAATACA	ATGAGCCCAA	GACTAACCGA	TCCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGAAT	TCCCAATATC	ATCTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACCTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACGT	TTCTTAGATC	TCCACATATG	1800
	AACCTGTGCG	GACTGTGAGA	ATCTTAAAT	ACAGTTTCTA	TAAACAGATA	TGAGGAGGAG	1860
85	AGTTTATGGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTCTGT	CATCTCCATT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAGACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAAAGAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100

	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCAAG	2220
	TCCTTTTCTG	CAGGCCCACT	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
5	CATTATTCTA	CCTTTGCCCTA	CTTCCCAACT	GAGGTAAAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACG	2400
	GTATACAATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAAGGC	AGTTATACCC	CTTGTGATCG	TGTCAGCCCT	GACTTTTATC	2520
	TGCTAGTAGG	TTCTTGTGGG	TATTCTCATC	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
10	TTTTACTTAG	AGGACAGTAC	ATCCCTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTT	2640
	CCAATTTTCA	ATGATGTGGG	AGCAATTCCA	ATAAAGCACT	TTCCAAAGCA	TGTTGCAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTGAAGAA	TTTGAGACAC	TGAAGAGTTT	TTACCAAGAA	2760
	GTGCAGAGCT	GTACTGTTGA	CTTAGGTATT	ACAGCAGACA	GCTCCAACCA	CCCAGACAAC	2820
	AAGCACAAGA	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAAGCTAGCA	2880
	CAGCTTGCTG	AAAAGGATGG	CAAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
15	TACAACAGAC	CAAAAGCTTA	TATTGCTGCC	CAAGGCCAC	TGAATCCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAGTTAATT	TCATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAAGTG	TGATCAGTAC	TGGCCTGGCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACTTTCTGG	TCACTCAGAA	GAGTGTGCAA	GTGCTTGCC	ATTATACTGT	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAAATT	AAAAAAGGCG	TCCAGAAAG	GAAGACCCAG	TGGACGTGTG	3240
20	GTCAACAGT	ATCACTACAC	GCAGTGGCCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGTGAGAAA	GGCAGCCTAT	GCCAGCGCC	ATGCAGTGGG	GCCTGTTGTC	3360
	GTCCACTGCA	TGCTGGAGT	TGGAAGAACA	GGCACAATA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTC	AACAGSAGG	AACGTGCAAC	ATATTGGCT	TCTTAAACAA	CATCCGTCCA	3480
	CAAGAAATTT	ATTTGGTACA	AACGTAGGAG	CAATATGTCT	TCATTTCATG	TACACTGTTT	3540
25	GAGGCCATAC	TTAGTAAAGA	AACGTAGGAG	CTGGACAGTC	ATATTTCATG	CTATGTTAAT	3600
	GCACCTCTCA	TTCTGGGACC	AGCAGGCAAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGAATAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAATCGAA	CTTCTCTTAT	CATCCCTGTG	GAAAGATCAA	GGGTGGCAT	TTTATCCCTG	3780
	AGTGAGAGAG	GCACAGACTA	CATCAATGCC	TCTATATATCA	TGGGCTATTA	CCAGAGCAAT	3840
30	GAATTCATCA	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTTCTG	GAGGATGATA	3900
	TGGGACCATTA	ATGCCCAACT	GGTGGTTATG	ATTCCTGATG	GCCAAAACAT	GGCAGAAAGT	3960
	GAATTTGTTT	ACTGGCCAG	TAAAGATGAG	CCTATAAATT	GTGAGAGCTT	TAAGGTCACT	4020
	CTTATGSGCTG	AGAACAACAA	ATGTCTATCT	AATGAGGAAA	AACCTATAAT	TCAGGACTTT	4080
	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTTGAAGTGA	GGCACTTTCA	GTGCTCTAAA	4140
35	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAGAA	4200
	GAGCTGCGCA	ATAGGGATGG	GCCTATGATT	GTTCATGATG	AGCATGGAGG	AGTGACGGCA	4260
	GGAACTTTCT	GTGCTCTGAC	AACCCCTATG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCAGG	TAGCCAGAT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATTGAG	4380
40	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTGA	GCACAGGCA	GGAAGAGAAT	4440
	CCATCCACCT	CTCTGGACAG	TAATGGTGCA	GCATTGCGCT	ATGGAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAAC	CAGAAAGGGG	TGGGGGGACT	CACATCTGAG	CATTGTCTTC	4560
	CTCTCTCTAA	AAITAGGCG	GAATATCAGT	CTAGTTCTGT	TATCTGTGTA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTTCATGA	CATAGGATTC	TGCCGCCAAA	TTTATATCAT	TAACAATGTG	4680
45	TGCCCTTTTG	CAAGACTTGT	AATTTACTTA	TTATGTTTGA	ACTAAAATGA	TGGAATTTTA	4740
	CAGTATTTCT	AAGAATGGAA	TTGTGGTATT	TTTTCTGTA	TTGATTTTAA	CAGAAAATTT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAAACTA	CAGAAAATGT	TTGTTTTTAG	TGTCAAATTT	4860
	TTAGCTGTAT	TTGTGCAAT	TATCAGGTTT	GCTAGAAATA	TAACTTTTAA	TACAGTAGCC	4920
	TGTAATAATA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCACT	ATTCACCTAA	4980
50	AGTAGAATA	ATCTGTTACT	TATTTGTAAT	ACTGCCCTAG	TGCTCTCATG	GACCAAAATTT	5040
	ATATTATATA	TTGTTTATTA	TTATATTTTA	CTACTGAGTC	AAGTTTCTTA	GTCTGTGTA	5100
	ATTGTTTAGT	TTAATGACGT	AGTTTATTAG	CTGCTCTTAC	TCTACCACTT	TTCTGACATT	5160
	GTATTGTGTT	ACCTAATGCT	TTAATTTTGT	TTTACGATGT	AAITTTAACT	TTTGTGGAAA	5220
	ATAGAAATAC	CTTCAATTTG	AAAGAAGTTT	TTATGAGAAT	AACACCTTAC	CAACATTTGT	5280
55	TCAAATGTTT	TTTATCCAA	GAATTGCAAA	AATAAATATA	AATATTGCCA	TTAAAAAAA	5340
	AAAAAAA	AAAAAAA	AAAAAAA				

Seq ID NO: 575 Protein sequence:
Protein Accession #: Eos sequence

60	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQRNNG	KKYPTCNPSK	60
	QSPINIDEDL	TQVNVNLKLL	KPQGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
65	FKASKITFHW	GKCNMSDDGS	EHSLEGQKFP	LEMQIYCFDA	DRPSSFEAAV	KGKGKLRALS	180
	ILFEVGTGTEEN	LDFKALIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYVI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVPCEVLTMQ	QSGYVLMMDY	LQNNFRBQQY	KFSRQVPSY	300
	TGKEEIHFAV	CSSEPENVOA	DPENYTSLLV	TWERPRVVDY	TMIEKFAVLY	QQLDGEDQTK	360
70	HEFLTDGVDQ	LGAILNNLLP	NMSYVLQIVA	ICTNGLYKRY	SDQLIVDMPT	DNPELDLPPE	420
	LIGTEEIIKE	EEEGKIDIEG	AIVNPGRDSA	TNQIRKKEPO	ISTTHYNRI	GTKYNEAKTN	480
	RSPTRGSEFS	GKADVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPFET	VEGTSASLND	540
	GSKTVLRSPH	MNLSGTAESL	NTVSIETEEY	ESLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPETITVD	VLIPESARNA	SEDSTSSGSE	ESLKDPMSBG	NWVFPSTDI	660
75	TAQPPDVSGSR	ESFLQNTYTE	IRVDESEKTT	KSPSAGPVMS	QGPSVTDLEM	PHYSTPAYFP	720
	TEVTTHAFTP	SSRQDLVST	VNVVYSQTTQ	PVYNAAEASNS	SHERIGLAE	GLESEKKAVI	780
	PLVIVSALTFF	ICLVVLVGLL	IYWRKCFQTA	HFYLEDSTSP	RVISTPTPTI	FPISDDVGAI	840
	PIKHFPKHVA	DLHASSGFTE	EPETLKRFYQ	EVQSCVTDLG	ITADSSNHDP	NKHKMRYINI	900
	VAYDHSRVKL	AQLAEKDGKL	TDYINANYVD	GYNRPKAYLA	AQGLPKSTAE	DFWRMIWEHN	960
80	VEVIVMITNL	VEKGRKRCQD	YWPADGSEY	GNFLVTQKSV	QVLAYYTVRN	FTLRNTRIKK	1020
	GSQKGRPSGR	VVTQVHYTQ	PDMGVPEYSL	PVLTFFVRKAA	YAKRHAVGVP	VVHCSAGVGR	1080
	TGTYIVLDSM	LQIQHEGTV	NIPGFLKHIR	SORNYLVQTE	EQYVFIHDTL	VEAILSKETE	1140
	VLDSDIHAYV	NALLIPGPAG	KTKLEKQFQL	LSQSNIQQSD	YSAALKQCNR	EKNRTSSIIPI	1200
	VERSRVGISS	LSGEGTDYIN	ASYIMGYYS	NEPIITQHEPL	LHTIKDPWRM	IWDHNAQLVV	1260
85	MIPDQGNMAE	DEFVYVWPNK	EPINCESFKV	TLMAEEHKCL	SNEEKLIQD	FILEATODDY	1320
	VLEVRHFPQCP	KWNPDPSPIS	KTFELISVIK	EAAANRDGPM	IVDEHGGVT	AGTFALCTTL	1380
	MHOLEKENSV	DVYQVAKMIN	LMRPGVFADI	EQYQFLYKVI	LSLVSTRQEB	NPSTSLDSNG	1440
	AALPDGNIAE	SLESIV					

Seq ID NO: 576 DNA sequence
Nucleic Acid Accession #: EOS sequence
Coding sequence: 148-4494

5	1	11	21	31	41	51	
	CACACATACG	CACGACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
10	CGCGAGGGGG	CCGACAGCCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCTATACAC	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
15	AACACATTCA	TTTATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTACGCGGAG	GAGTTTCAGA	AATGGTGTTT	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCACATT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCAGAG	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAAGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTG	720
20	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTAGTGC	GTTTTGAGGA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAACTCTAA	CTGACAACTA	TTACATTTC	840
	AATGGCTCAT	TGCAATCTCT	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTTTGTT	AAGTCTTAC	AATGCAACAA	960
	TCGTGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTCCAGAGCA	ACAGTACAAG	1020
25	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
	AGTTACAGAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAAAT	ATACCAAGCT	TCTTGTTACA	1140
	TGGGAAAGAC	TCGAGTCTGT	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTTGTACCGA	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTTGA	CAGATGGCTA	TCAGACTTGT	1260
	GGTGCTATTCT	TCATAAATTT	GCTACCCAAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
30	TGCATAATG	CGTTTATATG	AAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGCT	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGAA	AGACATTTGA	AGAAGGCGCT	ATTGTGAATC	CTGCTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAGGA	ACCCAGATTT	TCTACCACAA	CACACTACAA	TCCGATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
35	AAGGGTGAAG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAG	ATATTCTCTT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAACTG	TTCTTAGATC	TCCCATATG	1800
	AACCTGTGGG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTTA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
40	GCAACTTCTG	CTATCCCATT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTCTCTCC	1980
	GAAAACCCAG	AGACATTAAC	ATATGATGTC	CTTATACCAAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTACTCAGG	TTTCAAGGAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
	GTGTGTTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACGTGAGTA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
45	TCCTTTTCTG	CAGGCCAGTG	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTGGCCTA	CTTCCCAACT	GAGGTAAAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTGGGT	CTCCACGCTC	AACTGGTAT	ACTGCGAGAC	AACCCAAACG	2400
	GTATACAAAT	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
50	CTAGTGGTTC	TTGTGGGAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
	ATTTCCAGAT	ATGTCGAGC	AATCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	TGGGGTTTAC	TGAAGAATTT	GAGGAAGTGC	AGAGCTGTAC	TGTTGACTTA	2760
	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAACAAGC	ACAAGAAATG	ATACATAAAT	2820
55	ATCGTTGCTC	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCAAA	2880
	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
	GCTGCCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTTC	GGAGAATGAT	ATGGGAACAT	3000
	AATGTGGAAG	TTATTGTGAT	GATAACAAAC	CTCGTGGAGA	AAGGAAGGAG	AAAATGTGAT	3060
	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAACT	TTCTGGTCTC	TCAGAAAGAT	3120
60	GTGCAAGTGC	TTGCTTATTA	TACTGTGAGG	AAATTTTACT	TAAGAACAAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGG	CGTGTGGTCA	CACAGTATCA	CTACACGAG	3240
	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGGCA	3300
	GCCTATGCCA	AGGCCATGTC	AGTGGGGCCT	GTGTGCTGCC	ACTGCACTGC	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAACT	3420
65	GTCAACATAT	TTGGCTTCTT	AAACACATC	CGTTCAAAA	GAAATTATTT	GGTACAAACT	3480
	GAGGAGCAAT	ATGTCCTCAT	TCATGATACA	CTGGTTGAGG	CCATACCTAG	TAAAGAAACT	3540
	GAGGTGCTGG	ACAGTCAAT	TCATGCCTAT	GTTAATGCAC	TCTTCATTCC	TGGACAGCA	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCAG	CTCCTGAGCC	AGTCAATAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAGAA	ATCGAACTTC	TTCTATCATC	3720
70	CCTGTGGAAA	GATCAAGGGT	TGGCATTTCA	TCCCTGAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGCCCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCATTAC	CCAGCACCCT	3840
	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	ACCATAATGC	CCAAGTGGTG	3900
	GTTATGATTCT	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACCTCTA	TGGCTGAAGA	ACACAAATGT	4020
75	CTATCTAATG	AGGAAAAACT	TATAATTTCAG	GACTTTATCT	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	CAAAATCCAGA	TAGCCCCATT	4140
	AGTAAACATT	TTGAACCTAT	AAGTGTATTA	AAAGAAGAAG	CTGCCAATAG	GGATGGGCTC	4200
	ATGATTGTTCT	TGATGAGACA	TGGAGGAGTG	ACGGCAGGAA	CTTTCTGTGC	TCTGACAACC	4260
	CTTATGCACC	AACATAGAAA	AGAAAATTC	GTGGATGTTT	ACCAGGTAGC	CAAGATGATC	4320
80	AATCTGATGA	GGCCAGGAGT	CTTGTCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACCTCTCT	GGACAGTAAT	4440
	GGTGACAGAT	TGCTGATGG	AAATATAGCT	GAGAGCTTAG	AGTCTTATGT	TTAACACAGA	4500
	AAGGGGTGGG	GGGACTCACA	CTGAGCATT	GTTTTCTCT	TCCTAAATTT	AGGCAGGAAA	4560
	ATCAGTCTAG	TTCTGTATTC	TGTTGATTTC	CCATCACCTG	ACAGTAACCT	TCATGACATA	4620
85	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	TTTTTGCAAG	ACTTGTAATT	4680
	TACTTATTAT	GTTTGAACATA	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
	GGTATTTTTT	TCTGTATTGA	TTTAAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTCC	4800

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

```

AACTACAGA AAATGTTTGT TTTTAGTGTG AAATTTTGTG CTGTATTGTG AGCAATTATC 4860
AGGTTTGTCTA GAAATATAAC TTTTAATACA GTAGCCTGTA AATAAACACAC TCCTCCATAT 4920
GATATTCAAC ATTTTACAAC TGCAATGATC ACCTAAAGTA GAAATAATCT GTTACTTATT 4980
GTAAATACTG CCCTAGTGTG TCCATGGACC AAATTTATAT TTATAATTGT AGATTTTAT 5040
ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTGTAATTG TTTAGTTTAA TGACSTAGTT 5100
CATTAGCTGG TCTTACTCTA CCAGTTTCTT GACATTGTAT TGTGTTACCT AAGTCATTAA 5160
CTTTGTTTCA GCATGTAATT TTAACTTTGG TGGAAATAG AAATACCTTC ATTTTGAAAG 5220
AAGTTTATAT GAGAAATAACA CCTTACCAA CATTTGTTCAA ATGTTTTTTA TCCAAGGAAT 5280
TGCAAAATAA AATATAAATA TTGCCATTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 5340
AAA
  
```

Seq ID NO: 577 Protein sequence:
Protein Accession #: EOS sequence

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

```

1 11 21 31 41 51
MRILKRFLAC IQLLCVCRLD WANGYYRQOR KLVEEIGWSY TGALEKQNWG KKYPCTNSPK 60
QSPINIDEDL TOVNVNLKKL KFGQWDTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120
FKASKITFWH GKNCMSDDGS EHSLEGQKFP LEMQIYCPDA DRFSSFEAV KKGKLRALS 180
ILFEVGTZEN LDPKAIIDGV ESVSFRGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240
TDTVDWIVFK DTVSISESQL AVFCEVLTMO QSGYVLMMDY LQNNFREQQY KFSRQVFSSY 300
TGKEEIHBAV CSSEPEVUQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
HEPLTDGYQD LGAILNLLP NMSYVLQIVA ICTNGLYGYK SDQLIVDMPT DNPELDLPPE 420
LIQTEBIEKE EEEGKDIEEG AIVNPGRDSA TNQIRKKBPQ ISTTTHYRNI GTKYNEAKTN 480
RSPTRGSEFS GKGDVNTSL NSTSQPVTKL ATEKDISLTS QVTTELPPHT VEGTSASLND 540
GSKTVLRSPH MNLSGTAEYL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600
ENISQGYIFS SENPETITDY VLIPESARNA SEDSTSSGSE ESLKDPMEG NVWFPSTDI 660
TAQPDVGSGR ESPLQNYTE IRVDESEKTT KSFSAGPVMS QGPSVDLEM PHYSTFAYFP 720
TEVTPHAPT SRQQLDVST VNVVYSQTTQ PVYNEASNS HESRIGLAEG LESEKKAVIP 780
LVIVSALTFI CLVVLVGLIL WYRKCFQTAH FYLEDSTSPR VISTPTPTIP PISDDVGAIP 840
IKHFPKHVAD LHASSGFTTE FEEVQSCTVD LGITADSSNH PDNKHKNRYI NIVAYDHSRV 900
KLAQLAEKDG KLTIDYINANY DGYNRPKAY IAAQGPLKST AEDFWRMINE HNVEIVMIT 960
NLVEKGRKRC DQYWPADGSE EYGNFLVTOK SVQVLAYTV RNFTLRNTKI KKGSKGRPS 1020
GRVVYQYHYT QWPDMDVPEY SLPVLTVFRK AAYAKRHAVG PVVHCASAGV GRTGTIYILD 1080
SMLQIQHEG TVNIPGFLKH IRSQRNYLVQ TEEQYVFIHD TLVEAILSKE TEVLDSHHA 1140
YVNAIILPBP AGTKLEKQF QLLSQSNIQO SDYSAALKQC NREKNRTSSI IPVERSRVGI 1200
SSLSGEGTDY INASYIMGYI QSNEFLITQH PLLHTIKDFW RMWDHNAQL VVMIPDGQNM 1260
AEDEFVYWFN KDEPINCESF KVTLMABEHK CLSNEKLII QDPFILEATQD DYVLEVRHPQ 1320
CPKWPNPDSF ISKTFELISV IKEEAAARND FMIVHDEHGG VTAGTFCALT TLMHQLEKEN 1380
SVDVYQVAKM INLMRPGVFA DIEQYQFLYK VILSLVSTRQ EENPSTSLDS NGAALPDGNI 1440
AESLESLV
  
```

Seq ID NO: 578 DNA sequence
Nucleic Acid Accession #: EOS sequence
Coding sequence: 501-4514

50
55
60
65
70
75
80
85

```

1 11 21 31 41 51
CACACATACG CACGCACGAT CTCACCTGGA TCTATACACT GGAGGATTAA AACAAACAAA 60
CAAAAAAACC ATTTCTCTCG CTCCTCCCTCC CTCCTCCACTC TGAGAAGCAG AGGAGCCGCA 120
CGGCGAGGGG CCGCAGACCG CTTGAAATG CGAATCCTAA AGCGTTCTCT CGCTTGCAAT 180
CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
CTTGTGTAGG AGATTGGCTG GTCTATATCA GGAGCACTGA ATCAAAAAAT TGGGGAAAGA 300
AATATCCAACT ATGTAATAGC CCAAAACAA CTCTATCAAA TATTGATGAA GATCTTACAC 360
AAGTAAATGT GAATCTTAAG AAACCTAAAT TTCAGGGTTG GGATAAAACA TCATTGGAAA 420
ACACATTCAT TCATAACACT GGGAAACAG TGGAAATTAA TCTCACTAAT GACTACCGTG 480
TCAGCGGAGG AGTTTCAGAA ATGGTGTGTA AAGCAAGCAA GATAACTTTT CACTGGGGAA 540
AATGCAATAT GTCATCTGAT GGATCAGAGC ATAGTTTAGA AGGACAAAAA TTTCCACTTG 600
AGATGCAAT CTACTGCTTT GATGCGGACC GATTTCGAAG TTTTGAGGAA GCAGTCAAG 660
GAAAGGGGAA GTTAGAGACT TTATCCATT TGTGTGAGGT TGGGACAGAA GAAATTTGG 720
ATTTCAAAGC GATTATTGAT GGAGTCGAAA GTGTTAGTCC TTTTGGGAAG CAGGCTGCTT 780
TAGATCCATT CATACTGTGT AACCTTCTGC CAAACTCAAC TGACAAGTAT TACATTACA 840
ATGGCTCATT GACATCTCCT CCTGACAGC ACACAGTTGA CTGGATTGTT TTTAAAGATA 900
CAGTTAGCAT CTCTGAAAGC CAGTTGGCTG TTTTGTGTA AGTTCTTACA ATGCAACAAT 960
CTGGTTATGT CATGCTGATG GACTACTTAC AAAACAATTT TCGAGAGCAA CAGTACAAGT 1020
TCTCTAGACA GGTGTTTTCC TCATACACTG GAAAGGAAGA GATTATGAA GCAGTTTGTA 1080
GTTCAGAAC AGAAATATGT CAGGCTGACC CAGAGAAATTA TACCAGCCTT CTTGTTACAT 1140
GGGAAAGACC TCGATCGTT TATGATACCA TGATTGAGAA GTTTGCAATT TTGTACCAGC 1200
AGTTGGATGG AGAGGACCAA ACCAAGCATG AATTTTGAC AGATGGCTAT CAAGACTTGG 1260
GTGCTATTCT CAATAATTGT CTACCCAATA TGAGTTATGT TCTTCAGATA GTAGCCATAT 1320
GCACATAAGG CTTATATGGA AATACAGCG ACCAAGTATG TGTCCGATG CCTACTGATA 1380
ATCCTGAATG TGATCTTTTC CTTGAATTAA TTGGAAGTGA AGAAATAATC AAGGAGGAGG 1440
AAGAGGGGAA AGACATTGAA GAAGGCGCTA TTGTGAATCC TGGTAGAGAC AGTGCTACAA 1500
ACCAAAATCAG GAAAGAGGAA CCCCAGATT CTACCAACAC ACTACATAT CGCATAGGGA 1560
CGAAATACAA TGAAGCCAAG ACTAACCGAT CCCCACAAAG AGGAAGTGAA TTCTCTGGAA 1620
AGGGTGATGT TCCCAATACA TCTTTAAATT CCACTTCCCA ACCAGTCACT AAATTAGCCA 1680
CAGAAAAAGA TATTTCTTTG ACTTCTCAGA CTGTGACTGA ACTGCCACTC CACACTGTGG 1740
AAGGTACTTC AGCCTCTTTA AATGATGGCT CTAAAAGTGT TCTTAGATCT CCACATATGA 1800
ACTTGTGGGG GAGCTGAGAA TCCTTAAATA CAGTTCTAT AACAGAAAT GAGGAGGAGA 1860
GTTTATTGAC CAGTTTCAAG CTGTACTGAG GAGCTGAAGA TTCTTCAGGC TCCAGTCCCG 1920
CAACTTCTGC TATCCCATTC ATCTCTGAGA ACATATCCCA AGGGATATATA TTTCTCTCG 1980
AAAAACCAAG GACATAACA TATGATGTC TTATACCAAG ATCTGCTAGA AATGCTTCG 2040
AAGATTCAAC TTCACTCAGT TCAGAAAGAT CACTAAAGGA TCCTTCTATG GAGGGAATG 2100
TGTGTTTCC TAGCTTACA GACATAACAG CACAGCCCGA TGTGATGAT GGCAGAGAGA 2160
GCTTCTTCCA GACTAATTAC ACTGAGATAC GTGTTGATGA ATCTGAGAG AACAACCAAG 2220
CCTTTTCTGC AGGCCAGTG ATGTCACAGG GTCCCTCAGT TACAGATCTG GAAATGCCAC 2280
  
```

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85

ATTATTCTAC CTTTGCCCTAC TTCCCAACTG AGGTAACACC TCATGCTTTT ACCCCATCCT 2340
 CCAGACAAAC GAGTTTGGTC TCCACGGTCA ACGTGGTATA CTGCGAGACA ACCCAACCGG 2400
 TATACAATGA GGCCAGTAAT AGTAGCCATG AGTCTCGTAT TGGTCTAGCT GAGGGGTGG 2460
 AATCGAGAA GAAGCGAGTT ATACCCCTTG TGATCGTGTG AGCCCTGACT TTTATCTGTC 2520
 TAGTGGTCTT TGTGGGTATT CTCATCTACT GGAGGAAATG CTTCAGACT GCACACTTTT 2580
 ACTTAGAGGA CAGTACATCC CCTAGAGTTA TATCCACACC TCCAACACCT ATCTTTCCAA 2640
 TTTAGATGA TGTCCGAGCA ATTCCAATAA AGCACTTTCC AAAGCATGTT GCAGATTTC 2700
 ATGCAAGTAG TGGGTTTACT GAAGAATTG AGACACTGAA AGAGTTTAC CAGGAAGTGC 2760
 AGAGCTGTAC TGTGTACTTA GGTATTACAG CAGACAGCTC CAACCAACCA GACAACAAGC 2820
 ACAAGAATCG ATACATAAAT ATGCTGCTT ATGATCATAG CAGGTTAAAG CTAGCACAGC 2880
 TTGCTGAAAA GGATGGCAAA CTGACTGATT ATATCAATGC CAATTATGTT GATGGCTACA 2940
 ACAGACCAAA AGCTTATATT GCTGCCAAG GCCCACTGAA ATCCACAGCT GAAGATTTC 3000
 GGAGAAATGAT ATGGGAACAT AATGTGGAAG TTATTGTCAT GATAACAAAC CTCGTGGAGA 3060
 AAGGAAGGAG AAAATGTGAT CAGTACTGGC CTGCCGATGG GAGTGAGGAG TACGGGAAC 3120
 TTCTGGTCAC TCAGAAGAGT GTGCAAGTGC TTGCCTATTA TACTGTGAGG AATTTTACTC 3180
 TAAGAAACAC AAAAATAAAA AAGGGCTCCC AGAAAGGAAG ACCCAGTGA CGTGTGGTCA 3240
 CACAGTATCA CTACACGAGC TGGCCTGACA TGGGAGTACC AGAGTACTCC CTGCCAGTGC 3300
 TGACCTTGTG GAGAAGGCA GCCTATGCCA AGCGCCATGC AGTGGGGCCT GTTGTGCTCC 3360
 ACTGCAGTGC TGGAGTTGA AGAACAGGCA CATATATTGT GCTAGACAGT ATGTTGCAGC 3420
 AGATTCAACA CGAAGGAATC GTCAACATAT TTGGCTTCTT AAAACACATC CGTTCACAAA 3480
 GAAATTTATT GGTACAAACT GAGGAGCAAT ATGTCTTCAT TCATGATACA CTGTTGAGG 3540
 CCATACCTAG TAAAGAAACT GAGGTGCTGG ACAGTCAAT TCATGCTTAT GTTAATGCA 3600
 TCCTCATTC TGGACCAACA GGCACAAACA AGCTAGAGAA ACAATCCAG CTCCTGAGCC 3660
 AGTCAAAAT ACACAGAGT GACTATTCTO CAGCCCTAAA GCAATGCAAC AGGGAAGA 3720
 ATCGAATCTT CTCTATCATC CCTGTGGAAG GATCAAGGT TGGCATTTC TCCCTGAGTG 3780
 GAGAAGGCAC AGACTACATC AATGCCCTCT ATATCATGGG CTATTACAG AGCAATGAAT 3840
 TCATCATTC CAGACACCTC CTCTTCATA CCATCAAGGA TTTCTGGAGG ATGATATGGG 3900
 ACCATAATGC CCACTGTGTG GTTATGATTC CTGATGGCCA AAACATGGCA GAAGATGAAT 3960
 TTGTTTACTG GCCAAATAAA GATGAGCCTA TAAATGTGA GAGCTTTAAG GTCACCTTA 4020
 TGGCTGAAGA ACACAAATGT CTATCTAATG AGGAAAAACT TATAATTTCAG GACTTTATCT 4080
 TAGAAGCTAC ACAGGATGAT TATGTACTTG AAGTGAGGCA CTTCAGTGT CCTAAATGGC 4140
 CAAATCCAGA TAGCCCACTT AGTAAACTT TTGAACCTAT AAGTGTATA AAAGAAGA 4200
 CTGCCAATAG GGATGGGCTC ATGATTGTTT ATGATGAGCA TGGAGGAGTG ACGGAGGAA 4260
 CTTTCTGTGC TCTGCAACC CTTATGCACC AACTAGAAAA AGAAATTCG GTGGATGTT 4320
 ACCAGGTAGC CAAGATGATC AATCTGATGA GGCCAGGAGT CTTTGTGAC ATTGAGCAGT 4380
 ATCAGTTTCT CTACAAAGTG ATCCTCAGCC TTGTGAGCAC AAGGCAAGAA GAGAATCCAT 4440
 CCACCTCTCT GGACAGTAAT GGTGAGCAT TGCCTGATGG AAATATAGCT GAGAGCTTAG 4500
 AGTCTTTAGT TTAACACAGA AAGGGGTGGG GSGACTACA TCTGAGCATT GTTTCTCTCT 4560
 TCCTAAAATT AGGCAAGAAA ATCAGTCTAG TTCTGTTATC TGTGATTTT CCATCACCTG 4620
 ACAGTAACCT TCATGACATA GGATTCTGCC GCCAAATTTA TATCATTAAC AATGTGTGCC 4680
 TTTTTCAGAG ACTTGAATTT TACTTATTAT GTTTGAACCT AAATGATTGA ATTTTACAGT 4740
 ATTTCTAAGA ATGGAATGTG GTTATTTTTT TCTGTATTGA TTTTAACAGA AAATTTCAAT 4800
 TTATAGAGGT TAGGAATTC AAACACAGA AAATGTTTGT TTTTAGTGTC AAATTTTATG 4860
 CTGTATTGT AGCAATTATC AGGTTTGCTA GAAATATAAC TTTTAATACA GTAGCCTGTA 4920
 AATAAAACAC CTTTCCATAT GATATTCAAC ATTTTACAAC TGCAAGTATC ACCTAAGTA 4980
 GAAATAATCT GTTACTTATT GTAAATACTG CCCTAGTGTC TCCATGGACC AAATTTATAT 5040
 TTATAATCT AGATTTTTAT ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTGTAATG 5100
 TTTAGTTTAA TGACGTAGTT CATTAGCTGG TCTTACTCTA CCACTTTTCT GACATTGTAT 5160
 TGTGTACTCT AAGTCATTAA CTTTGTTCAT GCATGTAATT TTAACCTTTG TGGAAAAATG 5220
 AAATACCTTC ATTTTGAAGG AAGTTTTTAT GAGAATAACA CCTTACCAAA CATGTTTCAA 5280
 ATGGTTTTTA TCCAAGGAAT TGCAAAAAATA AATATAAATA TTGCCATTAA AAAAAAATA 5340
 AAAAAAATA AAAAAAATA AAA

Seq ID NO: 579 Protein sequence:
 Protein Accession #: EOS sequence

1 11 21 31 41 51
 MVFKASKITF HWGKCNMSSD GSEHSLEGQK PPLEMQIYCF DADRFSSFEE AVKKGKGLRA 60
 LSILFEVGTG ENLDPKAIID GVESVSRFGK QAALDPFILL NLLFNSTDRY YIYNGSLTSP 120
 PCTDVTVDIV FKDTVSISES QLAVFCEVLT MQQSGYVMLM DYLNQNFREQ QYKFSRQVFS 180
 SYTGKEEIEH AVCSSEPEPV QADPENYLSL LVTWERPRVV YDTMIEKFV LYQQLDGEDQ 240
 TRHEFLTDTG QDLGAILNLI LFNMSYVLQI VAICTNGLYG KYSDQLIVDM PTIDNPELDLP 300
 PELIGTEEII KEEEEKGDI EGAIVNPGRD SATNQIRKKE PQISTTTHYN RIGTKYNEAR 360
 TNRSPTRGSE FSGKQDVPNT SLNSTSQPVT KLATEKDI SL TSQTVELP HTVEGTSASL 420
 NDGSKTVLRS PEMNLSTAE SLNTVSTIEY EESLLTSFK LDTGAEDSSG SSPATSAIFP 480
 ISENISQGYI PSSENPETIT YDVLIPESAR NASEDSTSSG SEESLKDPSM EGNVWFPSST 540
 DITAGPDVGS GRESFLQNTY TEIRVDESEK TTKSFSAGPV MSQGPSVTDL EMPHYSTPAY 600
 FPTEVTPEAF TPSSRQDLV STVNVSYSQT TQPVYNEASN SSHESTRIGLA EGLESEKKAV 660
 IPLVIVSALT FICLVVLVGI LIYWRKCFQT AHFYLEDSTS PRVISTPTPT IFPISDDVGA 720
 IPIKHFPKHV ADLHASSGFT EEPETLKEFY QEVQSCVDL GITADSSNHP DNKHKNYIN 780
 IVAYDHSRVK LAQLAEKDGK LTDYINANYV DGYNRPKAYI AAQPLKSTA EDFWRMIWEH 840
 NVEIVMITN LVEKGRRKCD QYNPADGSEB YGNFLVTQKS VQVLAYYTVR NFTLRNTRIK 900
 KGSQKGRPSG RVVTQYHYTO WPDGMVPEYS LPVLTFVRKA AYAKRHAVGP VVVHCSAGVG 960
 RTGTIYVLDL MLQIQIHGEGT VNIFGFLKHI RSQRNYLVQT EEQYVFIHDT LVEAILSKET 1020
 EVLDSHIIHAY VNALLIPGPA GKTLEKQFQ LLSQSNIQQS DYSALKQCN REKNRTSSII 1080
 PVERSRVGIS SLSGEGTDYI NASYIMGYQ SNEFIITQHP LLHTIKDFWR MIWDHNAQLV 1140
 VMIPDGNMA EDEPYWPNK DEPINCESFK VTLMAHEHC LSNEEKLIQ DFLEATQDD 1200
 YVLEVRHFQC PKWPNPDSPI SKTFELISVI KEEAANDRGP MIVDEHGGV TAGTFCAALT 1260
 LMHOLEKENS VDVYQAKMI NLMRPGVFAD IEQYQFLYKV ILSLVSTRQB ENPSTSLDSN 1320
 GAALPDGNIA ESLESLV

Seq ID NO: 580 DNA sequence
 Nucleic Acid Accession #: EOS sequence
 Coding sequence: 148-4632

1 11 21 31 41 51

	CACACATACG	CACGACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCTCTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGCGAGGGGG	CGCGAGACCG	CTCGGAAATG	CGAATCTCTA	AACGTTTCCT	CGCTTGCAAT	180
5	CAGCTCCCTC	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCTATACA	GGAGCACTGA	ATCAAAAAAA	TGCGGGAAG	300
	AAATATCCAA	CATGTAAATG	CCCAAAACAA	TCTCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCAATTGGA	420
	AACACATCA	TTCAATAAC	TGGGAAACAA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
10	GTCAAGCGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTCTATCTG	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAGGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGAAG	GCAGGCTGCT	780
15	TTAGATCCAT	TCATCTGTTT	GAACCTTCTG	CCAACTCAA	CTGACAAGTA	TTACATTTC	840
	AATGGCTCAT	TGACATCTCC	TCCTGACACA	GACACAGTTG	ACTGGATTGT	TTTTAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTGTG	AAGTCTTAC	AATGCAACAA	960
	TCCTGGTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAG	1020
20	TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGAGT	TTTGTACAG	1200
	CAGTTGATGA	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTG	TCATAAATTT	GCTACCCCAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAAACAGC	GACCAACTGA	TTTTCGACAT	GCCTACTGAT	1380
25	AATCTCTGAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGT	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGCGCTT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATATC	GGAAAGGA	ACCCAGATT	TCTACCAACA	CACACTACAA	TGCGATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
30	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AAGTGCACCC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACCTG	TTCTTAGATC	TCCACATATG	1800
	AACCTGTGCG	GGACTGACGA	ATCTTAAAT	ACAGTTTCTA	TAAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTTGA	CCAGTTTCAA	GCTTGATCT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CACTCTGAG	AACATATCCC	AAGGTATAT	ATTTCTCTCC	1980
35	GAAGAACCCAG	AGACATAAAC	ATATGATGTC	CTTATACCAG	ATCTCTGATG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCTACAGG	TTCAGAAAG	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGTTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
	TCCTTTTCTG	CAGGCCCACT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
40	CATTATTCTA	CCTTTGCCCTA	CTTCCCAACT	GAGGTAAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACCAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACG	2400
	GTATACAAAT	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
	GAATCCGAGA	AGAAGCCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
45	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTTCCA	2640
	ATTTGAGATG	ATGTGCGAGC	AATTTCAATA	AAGCACTTTC	CAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTGA	CCAGGAAGTG	2760
	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	GCAGACAGCT	CCAACCACCC	AGACAACAG	2820
50	CACAAGAAATC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	2880
	CTTGTCTGAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	2940
	AACAGACCAA	AGGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTT	3000
	TGGAGAAATG	TATGGGAACA	TAAATGTGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	3060
	AAAGGAAGGA	GAAAAATGTA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	3120
55	TTTCTGGTCA	CTCAGAAAG	TGTGCAAGTG	CTTGCTTATT	ATACTGTGAG	GAATTTTACT	3180
	CTAAGAAATC	CAAAATATAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	3240
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	3300
	CTGACCTTTG	TGAGAAAGCC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTGCTC	3360
	CAGTGCATG	CTGGAGTTGG	AAGAACAGGC	ACATATATTT	TGCTAGACAG	TATGTTGCAG	3420
60	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGCTTCTT	TAAAAACAT	CCGTTACAAA	3480
	AGAAATTATT	TGCTACAAAC	TAGGAGGCAA	TATGTCTTCA	TTCATGATAC	ACTGGTTGAG	3540
	GGCATACTTA	GTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCATGCCTA	TGTTAATGCA	3600
	CTCCTCATTC	CTGACCCAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GGGTCTCACT	3660
	CTGTCAACCA	GGCTGGAGTG	CAGAGGCACA	ATCTCGGCTC	ACTGCAACCT	TCTCTCTCCT	3720
65	GGCTTAACCT	ATCCTCTTAC	CTCAGCTTCC	CGAGTGGCTG	GGACTATACT	CCTGAGCCAG	3780
	TCAAATATAC	AGCAGAGTGA	CTATTCTGCA	GGCTTAAAGC	AATGCAACAG	GGAAAGAAAT	3840
	CGAATCTCTT	CTATCATCCC	TGTGGAAGAA	TCAAGGGTTG	GCATTTTATC	CCTGAGTGGG	3900
	GAAGGCACAG	ACTACATCAA	TGCTCTCTAT	ATCATGGGCT	ATTACAGAG	CAATGAATTC	3960
	ATCATTACCC	AGCACCTCTT	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	4020
70	CATAATGCC	AACTGGTGGT	TATGATTCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	4080
	GTTTACTGGC	CAAAATAAAG	TGAGCCTATA	AATTGTGAGA	GCTTTAAGGT	CACCTTATAG	4140
	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTA	TAATTACAGG	CTTTATCTTA	4200
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCAT	TTCACTGTCC	TAAATGGCCA	4260
	AATCCAGATA	GGCCCACTAG	TAAACTTTT	GAACCTTATA	GTGTTTATAA	AGAAGAAGCT	4320
75	GGCAATAGGG	ATGGCCCTAT	GATTGTTTAT	GATGAGCATG	GAGGAGTGAC	GGCAGGAATC	4380
	TTCTGTGCTC	TGACAAACCT	TATGCACCAA	CTAGAAAAAG	AAAATTCCTG	GGATGTTTAC	4440
	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	4500
	CAGTTTCTCT	ACAAAGTGAT	CCTCAGCCTT	GTGGGCACAA	GGCAGGAAGA	GAATCCATCC	4560
	ACCTCTCTGG	ACAGTAAATG	TGCAGCATTG	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	4620
	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	4680
80	CTAAAAATTAG	GCAAGAAAAAT	CAGTCTAGTT	CTGTATCTG	TTGATTTCCC	ATCACCTGAC	4740
	AGTAACCTTC	ATGACATAGG	ATTCTGCCGC	CAAAATTATA	TCATTACCAA	TGTGTGCTTT	4800
	TTTGCAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACATAA	ATGATTGAAT	TTTACAGTAT	4860
	TTCTAAGAAAT	GGAAATGTGG	TATTTTTTTC	TGTAATTGAT	TTAACAGAAA	ATTTCAATTT	4920
	ATAGAGGTTA	GGAATTTCAA	ACTACAGAAA	ATGTTTGTTC	TTAGTGTCAA	ATTTTGTAGT	4980
85	GTAATTGTAG	CAATTATCAG	GTTTGCTAGA	AATATAACTT	TTAATACAGT	AGCCTGTAAA	5040
	TAAACACTTC	TTCCATATGA	TATTCAACAT	TTTACAACTG	CAGTATTTCAC	CTAAAGTAGA	5100
	AATAATCTGT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	5160

ATAATTGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTTCTG TGTAATTGTT 5220
 TAGTTTAAAG ACGTAGTTCA TTAGCTGGTC TTAGCTCTACC AGTTTTCTGA CATTGTATTG 5280
 TGTTAACTAA GTCAATTAAC TTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAATAGAA 5340
 ATACCTTCAT TTGAAAGAA GTTTTTATGA GAATAACACC TTACCAACA TGTTCAAAT 5400
 GGTTTTTATC CAAGGAATTG CAAAAATAA TATAATATT GCCATTAAAA AAAAAAATA 5460
 AAAAAAAAA AAAAAAAAA A

Seq ID NO: 581 Protein sequence:
 Protein Accession #: EOS sequence

1 11 21 31 41 51
 MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALEQKNWG KKYPTCNSPK 60
 QSPINIDEDL TQVNVNKKL KPQGWKTSK ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120
 FKASKITPHW GKNCMSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEAV KKGKLRALS 180
 ILFEVTEEN LDFKATIDGV ESVSRFGKQA ALDPFILLNL LFNSTDKYI YNGSLTSPPC 240
 TDTVDNIVFK DTVSISESQL AVFCEVLTMO QSGVVMMDY LQNNFRBQQY KFSRQVFSY 300
 TGKEEIEHAV CSSEPVNQV DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360
 HEFLTGYQD LGAILANLLP NMSYVLQIVA ICTNGLYGYK SDQLIVDMPT DNPEDLDFPE 420
 LIGTEEIIKE EEBGKIDEBG AIVNPRGDSA TNQIRKKEPQ ISTTHYNRI GTKYNEAKTN 480
 RSPTRGSEFS GKGDPVNTSL NSTSQPVTKL ATEKDISLTS QVTTELPPHT VEGTSASLND 540
 GSKTVLRSPH MNLSGTAESL NTVSITEYEE ESLTSFKLD TGAEDSSGSS PATAIPFIS 600
 ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSEMG NWVFPSTDI 660
 TAQFDVGSGR ESFLQNTYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM PHYSTFAYFP 720
 TEVTPHAFTP SSRQQLDVT VNVVYSQTTQ PVYNEASNSS HESRIGLAEG LESEKKAVIP 780
 LVIVSALTPI CLVVVLGILI YWRKCFQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840
 IKHFPKHVAD LHASSGFTEE FETLKEFYQE VQSCVDLGI TADSENHPDN KHNRYINIV 900
 AYDHSRVKLA QLAERKDKLT DYINANYVDG YNRPKAYIAA QGPLKSTAEF FWRMIWEHNV 960
 EVIVMTNLV EKRGRKCDQY WPAAGSEBYG NPLVTQKSVQ VLAYYTVRNP TLRNTKIKKG 1020
 SQKRPSPGRV VTQYHYTQWP DMGVPEYSLP VLTFRKAAY AKRHAAGPVV VHCAGVGR 1080
 GTYIVLDSML QQIQHEGTWN IFGLKHRS QRNYLVQTEE QYVPIHDTLV EAILSKEDEV 1140
 LDSHIHAYVN ALLIPGPAGK TKLEKQPQGL TSPRLCEGR TISAHCNPLP POLTDPPTSA 1200
 SRVAGTILIS QSNIQSDYS AALKQCNREK NRTSSIIPE RSRVGISLS EGEDYDINAS 1260
 YIMGYQSNF PIITQHPLH TIKDFWRMIV DHNAQLVMMI PDGQMAADE FVYWNKDEP 1320
 INCSEFKVTL MAEHLKLSN EEKLIQDFI LEATQDDYVL EVRHFQCPKW PNPDSPISK 1380
 FELISVKEE ANRNDGFMIV HDEHGGVTAG TFCALTLMH QLEKENSVDV YQVAKMINLM 1440
 RGVFADIEQ YQFLYKVLIS LVGTRQENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 582 DNA sequence
 Nucleic Acid Accession #: NM_002851.1
 Coding sequence: 148..7092

1 11 21 31 41 51
 CACACATAG CAGCAGCAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAATC ATTTCTCTCG CTCGCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCGCGA 120
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCTTAA AGCGTTTCCCT CGCTTGCAAT 180
 CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTGTGTGAAG AGATTGGCTG GTCCATACA GGAGCACTGA ATCAAAAAA TTGGGGAAG 300
 AAATATCCAA CATGTAATAG CCCAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAC ATCATTGGAA 420
 AACACATTC TCCATAACAC TGGGAAACAA GTGGAAATTA ATCTCACTAA TGACTACCGT 480
 GTCAGCGGAG GAGTTTCAGA AATGGTGT TTAAAGCAAGCA AGATAACTTT TCACCTGGGA 540
 AAATGCAATA TGTACTCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600
 GAGATGCAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAA 660
 GGAAGAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAAATTG 720
 GATTCAAGAG CGATTATTGA TGGAGTCGAA AGTGTGTAGT GTTTTGGGAA CGAGCTGTCT 780
 TTAGATCCAT TCATCTGTGT GAACCTTCTG CCAACTCAA CTGACAGATA TTACATTATC 840
 AATGGCTCAT TGCATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900
 ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTTTGTG AAGTCTTAC AATGCAACAA 960
 TCTGTTATG TCATGCTGAT GGACTACTTA CAAACAAATT TTGAGAGCA ACAGTACAAG 1020
 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAGGAAG AGATTCTATG AGCAGTTTGT 1080
 AGTTCAGAAC CAGAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTATCA 1140
 TGGGAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGAGT TTTGTACAG 1200
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260
 GGTGCTATTC TCAATAATTT GCTACCCAA ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTCGACAT GCTACTGTAT 1380
 AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAAGCT AAGAAATAAT CAAGGAGGAG 1440
 GAAGAGGGA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA 1500
 AACCAATCA GGAAGAGGA ACCCCAGATT TCTACCAACA CACACTACAA TCGCATAGGG 1560
 ACGAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620
 AAGGTTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680
 ACAGAAAGAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740
 GAAGTACTT CAGCTCTCTT AAATGATGGC TCTAAAACTG TTCTTAGATC TCCATATG 1800
 AACTTGTGCG GAGCTGCAGA ATCTTAAAT ACAGTTTCTA TAACAGAATA TGAGAGGAG 1860
 AGTTTATTGA CCACTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920
 GCAACTTCTG CATCTCCAT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980
 GAAACCCAG AGACAATAAC ATATGATGTC CTTATACAG AATCTGCTAG AAATGCTTCC 2040
 GAAGATTCAA CTTCATCAGG TTCAGAAGAA TCACATAAGG ATCTTCTAT GAGGAGAAAT 2100
 GTGTGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160
 AGCTTTCTCC AGACTAATA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220
 TCCTTTCTCG CAGGCCAGT GATGTACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA 2280
 CATATTCTA CTTTGCCTA CTTCCCACT GAGGTAAAC CTCTGCTTT TACCCCATCC 2340
 TCCAGACAAC AGGATTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCCAACG 2400
 GTATACAATG GTGAGACACC TCTCAACCT TCCTACAGTA GTGAAGTCTT TCCTCTAGTC 2460
 ACCCCTTGTG TGCTTGACAA TCAGATCCTC AACACTACCC CTGCTGCTTC AAGTAGTGAT 2520

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

TCGGCTTGC ATGCTACGCC TGTATTCCCT AGTGTGATG TGTCAATTGA ATCCATCCTG 2580
 TCTTCTATG ATGGTGCACC TTTGCTTCCA TTTTCTCTG CTCTCTCAG TAGTGAATTG 2640
 TTTGCGCATC TGCATACAGT TTCTCAAAAT CTCTCCACAAG TTACTTCCAG TACCGAGAGT 2700
 GATAAGGTGC CCTTGCATGC TTCTCTGCCA GTGGCTGGGG GTGATTGTCT ATTAGAGCCC 2760
 AGCCTTGCTC AGTATTCTGA TGTGCTGTCC ACTACTCATG CTGCTTCAGA GACGCTGGAA 2820
 TTTGGTAGTG AATCTGGTGT TCTTTATAAA ACGCTTATGT TTTCTCAAGT TGAACCAACC 2880
 AGCAGTATG CCATGATGCA TGCACGTTCT TCAGGGCCTG AACCTTCTTA TGCCTTGTCT 2940
 GATAATGAGG GCTCCCAACA CATCTTCACT GTTCTTACA GTTCTGCAAT ACCTGTGCAT 3000
 GATTCTGTGG GTGTAACCTA TCAGGGTTCC TTATTAGCG GCGCTAGCCA TATACCAATA 3060
 CCTAAGTCTT CGTTAATAAC CCCAAGTCCA TCATTACTGC AGCCTACTCA TGCCTCTCT 3120
 GGTGATGGGG AATGGTCTGG AGCCTCTTCT GATAGTGAAT TTCTTTTACC TGACACAGAT 3180
 GGGCTGACAG CCCTTAACAT TTCTTCACT GTTCTGTAG CTGAATTAC ATATACAACA 3240
 TCTGTGTTTG GTGATGATAA TAAGGCGCTT TCTAAAAGT AAATAATATA TGGAAATGAG 3300
 ACTGAAGTGC AAATCTCTTC TTCAATGAG ATGGTTTACC CTCTGAAAG CACAGTCATG 3360
 CCCAACATGT ATGATAATGT AAATAAGTTG AATGCGTCTT TACAAGAAAC CTCTGTTTCC 3420
 ATTTCTAGCA CCAAGGGCAT GTTTCACGGG TCCCTTGCTC ATACCACCAC TAAGGTTTTT 3480
 GATCATGAGA TTAGTCAAGT TCCAGAAAT AACTTTTCTG TTCAACCTAC ACATACTGTC 3540
 TCTCAAGCAT CTGGTGACAC TTGCTTAAA CCTGTGCTTA GTGCAAACT AGAGCCAGCA 3600
 TCCCTGAGCC CTGCTTCTAG TGAATGTTA TCTCCTTCAA CTCAGCTCTT ATTTTATGAG 3660
 ACCTCAGCTT CTTTTAGTAC TGAAGTATTG CTACAACCTT CAGCTTCAAGG TTCTGATGTT 3720
 GACACCTTGC TTTAAACTGT TCTTCCAGCT GTGCCAGTG ATCCAATATT GGTGTAAACC 3780
 CCCAAAGTTG ATAAATGAT TTCTACAATG TTGCATCTCA TTGATCAAA TTCTGCTTCA 3840
 AGTGAANAAC TGTGTCATCT TACATCTGTA CCAGTTTGTG ATGTGTGCGC TACTTCTCAT 3900
 ATGCACTCTG CTTCACCTCA AGGTTTGACC ATTTCTTATG CAAGTGAGAA ATATGAACCA 3960
 GTTTTGTAAA AAGTGAAGAG TTCCACCAA GTGGTACCTT CTTTGTACAG TAATGATGAG 4020
 TTGTTCCAAA CGGCCAATTT GGAGATTAAC CAGGCCCATC CCCCAAAAGG AAGGCATGTA 4080
 TTTGCTACAG CTGTTTATTC AATTGATGAA CCATTAAATA CACTAATAAA TAAGCTTATA 4140
 CATTCCGATG AAATTTTAACT CTCCACCAA AGTTCTGTTA CTGGTAAGGT ATTTGCTGGT 4200
 ATTTCAACAG TTGCTTCTGA TACATTGTGA TCTACTGATC ATTTCTGTTCC TATAGGAAAT 4260
 GGGCATGTGG CCATTACAGC TGTTTCTCCC CACAGAGATG GTTCTGTAACT CTCACAAAG 4320
 TTGCTGTTTC CTCTAAGGC AACTTCTGAG CTGAGTCATA GTGCCAAATC TGATGCCGGT 4380
 TTAGTGGTGG CTGGTGAAGA TGGTGACACT GATGATGATG GTGATGATGA TGATGACAGA 4440
 GATAGTGATG GCTTATCCAT TCATAAGTGT ATGTCTGCTC CATCTATAG AGAATCAGAG 4500
 GAAAAGGTAA TGAATGATTG AGACACCCAC GAAAACAGTC TTATGGATCA GAATAATCCA 4560
 ATCTCATACT CACTATCTGA GAATTCTGAA GAAGATAATA GAGTCACAAG TGTATCTCTA 4620
 GACAGTCAAA CTGGTATGGA CAGAAGTCTT GGTAAATCAC CATCAGCAAA TGGGCTATCC 4680
 CAAAAGCACA ATGATGGAAA AGAGGAAAT GACATTGAGA CTGGTATGTC TCTGCTTCTT 4740
 CTCAGCCCTG AATCTAAAGC ATGGGCAAGT CTGACAAAGT ATGAAGAAAG TGGATCAGGG 4800
 CAAGGTACTT CAGATAGCCT TAATGAGAAT GAGACTTCCA CAGATTTTCA TTTTGCAGAC 4860
 ACTAATGAAA AAGATGCTGA TGGGATCCTG GCAGCAGGTG ACTCAGAAAT AACTCCTGGA 4920
 TTCCACAGT CCCCAACATC ATCTGTACT AGCGAGAACT CAGAAGTGT CCACGTTTCA 4980
 GAGGAGAGG CCAATTAATG TAGCCATGAG TCTGCTATTG GTCTAGCTGA GGGGTGGAA 5040
 TCGAGAAGA AGGCAGTTAT ACCCCTTGTG ATCTGTGTCAG CCTGACTTT TATCTGTCTA 5100
 GTGGTCTCTG TGGGTATCTT CATCTACTGG AGGAAATGCT TCCAGACTGC ACACCTTTAC 5160
 TTAGAGGACA GTACATCCCT TAGAGTTATA TCCACACCTC CAACACCTAT CTTTCCAATT 5220
 TCAGATGATG TCGGAGCAAT TCCAATAAAG CACTTCCAA AGCATGTTGC AGATTTACAT 5280
 GCAAGTACTG GCTTACTGTA AGAATTGAG AACTGAAAG AGTTTACCA GGAAGTGCAG 5340
 AGCTGTACTG TTGACTTAGG TATTACAGCA GACAGCTCCA ACCACCCAGA CAACAGCAC 5400
 AAGAATCGAT ACATAAATAT CGTTGCTTAT GATCATAGCA GGGTTAAGCT AGCAGAGCTT 5460
 GCTGAAAAGG ATGGCAAACT GACTGATTAT ATCAATGCCA ATTATGTTGA TGGCTACAAC 5520
 AGACCAAAAG CTTATATTGC TGCCCAAGGC CCACTGAAAT CCACAGCTGA AGATTTCTGG 5580
 AGAATGATAT GGAACATAA TGTGGAAGTT ATTGTCTGTA TAACAACTT CGTGAGAAA 5640
 GGAAGGAGAA AATGTGATCA GTACTGGCCT GCCGATGGGA GTGAGGAGTA CGGGAACCTT 5700
 CTGGTCACTC AGAAGAGTGT GCAAGTGTCT GCCTATTATA CTGTGAGGAA TTTTACTCTA 5760
 AGAAACACAA AATAAATAAA GGGCTCCCAG AAAGGAAGAC CCACTGGAGC TGTGGTCACA 5820
 CAGTATCACT ACAGCGAGTG GCCTGACATG GAGGTACCA AGTACTCCCT GCCAGTGCTG 5880
 ACCTTTGTGA GAAAGGCAGC CTATGCCAAG CGCCATGCAG TGGGGCTGTG TGTCTCCAC 5940
 TGCAAGTCTG GAGTGGGAAG AACAGGCACA TATATTGTGC TAGACAGTAT GTTGACAGAG 6000
 ATTTCAACAG AAGGAAGTGT CAACATATTG GGCTTCTTAA AACACATCCG TTCACAAAGA 6060
 AATTATTGGG TACAAACTGA GGAGCAATAT GTCTTCATTC ATGATACACT GGTGTAGGCC 6120
 ATACTTAGTA AAGAACTGA GGTGCTGAC AGTCATATTC ATGCTATGT TAATGCACTC 6180
 CTCATTCTCT GACCAGCAGG CAAAACAAAG CTAGAGAAAC AATTCCAGCT CCTGAGCCAG 6240
 TCAAAATATC AGCAGAGTGA CTATTCTGCA GCCCTAAAGC AATGCAACAG GGAAGAAGAT 6300
 CGAACTTCTT CTATCATCCC TGTGGAAGA TCAAGGGTTG GCATTTATC CTGAGTGGGA 6360
 GAAGGCACAG ACTACATCAA TGCCCTCTAT ATCATGGGCT ATTACCAGAG CAATGAATTC 6420
 ATCATATCCC AGCAGCCTCT CCTTCATACC ATCAAGGATT TCTGGAGGAT GATATGGGAC 6480
 CATAATGCCC AACTGGTGGT TATGATTCTT GATGGCCAAA ACATGGCAGA AGATGAATTT 6540
 GTTTACTGGC CAAATAAAGA TGAGCCTATA AATTGTGAGA GCTTTAAGGT CACTCTTATG 6600
 GCTGAAGAAC ACAAATGTCT ATCTAATGAG GAAAACTTA TAATTCAAGG CTTTATCTTA 6660
 GAAGCTACAC AGGATGATTA TGTACTTGA GTGAGGCACT TTCAGTGTCC TAAATGGCCA 6720
 AATCCAGATA GCGCCATTAG TAAACCTTTT GAACCTTATA GTGTTATAAA AGAAGAGCT 6780
 GCCAATAGGG ATGGGCTTAT GATTGTTTAT GATGAGCATG GAGGAGTGAC GGCAGGAAC 6840
 TTCTGTGCTC TGACAACCTT TATGCACCAA CTAGAAAAAG AAAATTTCCG GGATGTTTAC 6900
 CAGGTAGCCA AGATGATCAA TCTGATGAGG CCAGGAGTCT TTGCTGACAT TGAGCAGTAT 6960
 CAGTTTCTCT ACAAGTGTAT CCTCAGCCTT GTGAGCACAA GGCAGGAAGA GAATCCATCC 7020
 ACCTCTCTGG ACAGTAATGG TGCAGCATTG CCTGATGAAA ATATAGCTGA GAGCTTAGAG 7080
 TCTTTAGTTT AACACAGAAA GGGGTGGGGG GACTCACATC TGAGCAITGT TTTCTCTTTC 7140
 CTAAAATTAG CAGGAAAAAT CAGTCTAGTT CTGTTATCTG TTGATTTCCC ATCACTTGAC 7200
 AGTAACCTTC ATGACATAGG ATCTGCGGCG CAAATTTATA TCATTACAA TGTGTGCTTC 7260
 TTTGCAAGAC TTGTAATTTT CTATTATGT TTGAACATAA ATGATTGAAT TTTACAGTAT 7320
 TTCTAAGAA GGAATGTGTG TATTTTTTTC TGTATTGATT TTAACAGAAA ATTTCAATTT 7380
 ATAGAGTTTA GGAATTTCAA ACTACAGAAA ATGTTTGTTC TTAGTGTCAA ATTTTATGCT 7440
 GTATTGTGTA CAATTATCAG GTTTGCTAGA AATATACTG TTAATACAGT AGCCTGTAAA 7500
 TAAACACTTC TTCCATATGA TATTCAACAT TTTCAACAGT CAGTATTAC CTAAAGTAGA 7560
 AATAATCTGT TACTTCTGGT AAATACTGCC CTAGTGCTC CATGGACCAA ATTTATATT 7620
 ATAATGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTCTG TGTAAATGTT 7680
 TAGTTAATG ACGTAGTTCA TTAGCTGGTC TTACTCTACC AGTTTCTGA CATTGTATTG 7740

TGTTACCTAA GTCATTAAC TTTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAAAATAGAA 7800
 ATACCTTCAT TTTGAAAGAA GTTTTATGA GAATAACACC TTACCAACA TTTGTTCAAAAT 7860
 GGTTTTATC CAAGGAATG CAAAAATAA TATAAATAT GCCATTAAAA AAAAAAATAA 7920
 AAAAAAATAA AAAAAAATAA A

5

Seq ID NO: 583 Protein sequence
 Protein Accession #: NP_002842.1

10 1 11 21 31 41 51
 | | | | | |
 MRILKRFLAC IQLLCVCRID WANGYYRQOR KLVZEIGWSY TGALNQKQWG KKYPTCNPK 60
 QSPINIDEDL TQVNVNLKRL KFGQWDKTSI ENTPIHNTGK TVEINLTNDY RVSGGVSEMV 120
 FKASKITFWH GKCNMSDGS EHSLEGQKFP LEMQIYCFDA DRFSSFEAV KKGKGLRALS 180
 ILFEVGTBEN LDFKAIIDGV ESVSRRFGKQA ALDPFILLNL LPNSTDKYI YNGSLTSPPC 240
 15 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVLMMDY LQNNFREQQY KFSRQVFSSY 300
 TGKEEIHBAV CSSEPEVQA DPENYTSLLV TWERPRVVDY TMIEKPAVLY QQLDGEDQTK 360
 HEFLTDGYOD LGAILNLNLP NMSYVLQIVA ICTNGLYGY SDQLIVDMPT DNPELDLPE 420
 LIGTEBIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKCEPQ ISTTTHYRI GTKYNEAKTN 480
 RSPTRGSEPS KGDVNPNTSL NSTSQPVTKL ATEKDISLTS QVTTELPHHT VEGTSASLND 540
 20 GSKTVLRSPH MNLSTGAESL NTVSITEYEE ESLTTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGYIFS SENPETITYD VLIPESARNA SEDSTSSGSE ESLKDPSEMG NVWFPSSTDI 660
 TAQPDVGSGR ESFLQNTYIE IRVDESEKIT KSPSAGPVMS QQPSVTDLEM PHYSTFYFPP 720
 TEVTPHAPTP SSRQQLVST VNVVYSQTQ PVYNGETPLQ PSYSSEVFPL VTPLLLDNQI 780
 LNTTPAASS DSALHATPVF PSVDVSFESI LSSYDGAFLP PFSSASFSS LFRHLHTVSQ 840
 25 ILPQVTSATE SDKVPLHASL PVAGDILLLE PSLAQYSDVL STTHAASETL EPGSESGLVLY 900
 KILMFQSEVP PSSDAMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG 960
 SLFSGSHIP IPKSLITPT ASLLQPTHAL SGDGEWSGAS SDSEFLLEPT DGLTALNISS 1020
 PVSVAEFTYT TSVFGDNKA LSKSEIIYGN ETELQIPSPN EMVYSESTV MFNMYDNVVK 1080
 LNASLQETS SVSISSTKGMFP GSLAHTTKV FDHEISQVPE NNFVSQPTHY VSQASGDTSL 1140
 30 KPVLANSSEP ASSDPASSEM LSPSTQLLFY ETSASFSTV LLQPSQFASD VDTLLKTVLP 1200
 AVPSDILVE TPVKDKISST MLHLIVNSA SSENMLHSTS VFVFDVSPTS HMHSASLQGL 1260
 TISYASEKIE PVLLKSESH QVPSLYSND ELFPQANLEI NQAHPPKGRH VFATPVLSID 1320
 EPLNTLNLKL IHSDEILTST KSSVTGKVA GIPTVASDTP VSTDHVSPIG NGHVAITAVS 1380
 PHRDGVSST KLFPFSKATS ELHSAKSDA GLVGGGEDGD TDDGDDDDDD RDSGLSIHK 1440
 35 CMSCSYRES KKVWMDSDT HENSLMDQNN PISYSLSENS EEDNRVTSVS SDSQTGMDRS 1500
 PGKSPSANGL SQKHNDGKEE NDIQTGSALL PLSPEKAWA VLTSDEESGS CQGTSDSLNE 1560
 NETSTDPSFA DTNEKADGI LAAGDSEITE GFPOSPTSSV TSENSEVFHV SEAEASHSSH 1620
 ESRIGLABGL BSEKXAVIPL VIVSALTFIC LVVLVGLIY WRKCFQTAHF VLEDSTSPRV 1680
 ISTPPTPIPP ISDDVGAIP KHFPKQVADL HASGFTTEF ETLKEFYQEV QSCTVDLGIT 1740
 40 ADSSNHPDNK HKRYRINIYA YDHSRVKLAQ LAEKDGLTD YINANYVDGY NRPKAYIAAQ 1800
 GPLKSTAEFD WRMIWEHNVE VIVMITNLVE KGRRKCDQYV PADGSEYEGN FLVTQKSVQV 1860
 LAYTYTRNLT LRNTKKKGS QKGRPSGRV TQYHYTQWPD MGVPEYSLPV LTFVRKAAAYA 1920
 KRHAVGPPVV HCSAGVGRTO TYIVLDSMLQ QIQHEGTVNI FGFLKHRSQ RNYLVQTEEQ 1980
 YVFIHDTLVB AILSKETEVL DSHIAYVNA LLIPGPAKGT KLEKQPQLLS QSNIQSDYS 2040
 45 AALKQCNREK NRTSSIIPVE RSRVGISSL GEGTDYINAS YIMGYQSNF FIITQHPLH 2100
 TIKDFWRMIV DHNAQLVVMV PDGQNMMADE FVYFENKDEP INCESFKVTL MAEEHKCLSN 2160
 EEKLIQDFI LEATQDDYVL EVRHFQCPKW PNPDSPIKLT FELISVIKE AANRDGPMIV 2220
 HDEHGGVTAG TPCALITLHM QLEKENSVDV YQVAKMINLM RPGVPADIEQ YQFLYKVLIS 2280
 LVSTRQEEMP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 584 DNA sequence
 Nucleic Acid Accession #: NM_005688.1
 Coding sequence: 126..4439

55 1 11 21 31 41 51
 | | | | | |
 CCGGGCAGGT GGCCTCATGCT CCGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCTTGGAGC 60
 AGGGGCGCAG GAATTCATGAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCCTCTCAG 120
 AGAAGATGAA GGTATATCAT ATAGGAAAG AGTATATCAT CCCCAGTCTCT GGTATAGAA 180
 60 GTGTAGGGA GAGAACACAGC ACTTCTGGGA CGCACAGAGA CGGTGAAGAT TCCAAAGTCA 240
 GGAGAACTCG ACCGTGGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCGA GCCGAGGGCC 300
 TCTCTCTTGA TGCCCTCCATG CATTCTCAGC TCAGAACTCT GGATGAGGAG CATCCCAAGG 360
 GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACACCAGC 420
 ACCCAGTGA CAATGCTGGG CTTTTTCTCT GTATGACTTT TTGTTGGCTT TCTTCTCTGG 480
 65 CCGGTGTGGC CCACAAGAAG GGGGAGCTCT CAATGGAAGA CGTGTGGTCT CTGTCCAAAGC 540
 ACCAGTCTTC TGACGTGAAC TGCAGAAGAC TAGAGAGACT GTGGCAAGAA GAGCTGAATG 600
 AAGTTGGGCG AGACGCTGCT TCCCTGCGAA GGGTTGTGTG GATCTTCTGC CGCACCAGGC 660
 TCATCTGTGC CATCGTGTGC CTGATGATCA CGCAGCTGGC TGGCTTCAGT GGACCAGCCT 720
 TCATGTGTA ACACCTCTTG GAGTATACCC AGGCAACAGA GTCTAACCTG CAGTACAGCT 780
 70 TGTTGTAGT GCTGGGCTTC CTCCTGACGG AAATCGTGGC GTCTTGGTGG CTGTGACTGA 840
 CTTGGGCATT GAATACCGA ACCGGTGTCC GCTTGGCGGG GGCCATCTTA ACCATGGCAT 900
 TTAAGAAGAT CCTTAAGTTA AAGAACATTA AAGAGAAATC CCTGGGTGAG CTCATCAACA 960
 TTTGCTCCAA CGATGGGCAG AGAATGTTTG AGGCAGCAGC CGTGGGCAGC CTGCTGGCTG 1020
 GAGGACCCGT TGTGTCCATG TTAGGCATGA TTTATAATGT AATTATCTCT GGACCAACAG 1080
 75 GCTTCTCTGG ATCAGCTGTT TTTATCTCTT TTTACCCAGC AATGATGTTT GCATCAGGGC 1140
 TCACAGCATA TTTCAAGAGA AAATGCGTGG CGGCCAAGCA TGAACGTGTC CAGAAGATGA 1200
 ATGAAGTTCT TACTTACATT AAATTTATCA AAATGTATGC CTGGGTCAAA GCATTTTCTC 1260
 AGAGTGTCCA AAAAATCCGC GAGGAGGAGC GTCCGATATT GGAAAAAGCC GGGTATCTCC 1320
 AGGGATATCAC TGTGGTGGT GCTCCCATG TGGTGGTGAT TGCCAGCGTG GTGACCTTCT 1380
 80 CTGTTCTATG GACCTGGGCT TTCGATCTGA CAGCAGCACA GGCTTTCACA GTGGTGACAG 1440
 TCTTCAATTC CATGACTTTT GCTTTGAAAG TAACACCGTT TTCAGTAAAG TCCTCTCTCAG 1500
 AAGCCTCAGT GGCTGTGAC AGATTTAAGA GTTTGTTTCT AATGGAAGAG GTTCACATGA 1560
 TAAAGACAA ACCAGCCAGT CCTCACATCA AGATAGAGAT GAAAAATGCC ACCTTGGCAT 1620
 GGAAGTCTCT CCATCTCAGT ATCCAGAACT CGCCCAAGCT GACCCCAAAA ATGAAAAAAG 1680
 85 ACAAGAGGGG TTCCAGGGGC AAGAAAGAGA AGGTGAGGCA GCTGCAGCGC ACTGAGCATC 1740
 AGGCGGTGCT GGCAGAGCAG AAAGGCCACC TCCTCTCTGA CAGTGACGAG CGGCCAGTCT 1800
 CCAAGAGGA AGAAGGCAAG CACATCCACC TGGGCCACCT GCGCTTACAG AGGACACTGC 1860

	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AACCTGGTTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCATT	TCAGCCATTT	TAGGCCAGAT	GACCGTCTTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACTCTT	GCTTATGTGG	CCGAGCAGGC	CTGGATCCTC	AATGCTACTC	2040
5	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTG	CCTGAGGCC	GACCTGGCCA	TTCTTCCCAG	CAGCGACCTG	ACGAGATTG	2160
	GAGAGCGAGG	AGCCAACTTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCGGGCCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTT	CAGTGCCTTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
10	TTGTTACCCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGATCTTC	ATGAAAGAGG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACATGATGA	TTTAAATGGT	GACTATGCTA	2460
	CCATTTTAA	TAACTGTTTG	CTGGGAGAGA	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCAAGTG	TTACAGAAAG	AAGTCAACAAG	ACAAGGGTCC	TAAAAACAGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCAAGTCC	CTGGTCACTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
15	TCCTGGTTAT	TATGGCCCTT	TTCATGCTGA	ATGTAGGCAG	CACCGCCTTC	AGCACCTGGT	2760
	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAAACACCAC	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTAGG	TGACAGCATC	AAGGACAATC	CTCATATGCA	GTAATATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCTCG	CTGATCTCTA	AAGCCATTGG	AGGAGTGTCT	TTTGTCAAGG	2940
	GCAGCTGAGC	AGCTTCTCTC	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTGAAGGCC	3000
20	CTATGAAGTT	TTTTGACAGC	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
	TGGATGAAGT	TGACGTGCGG	CTGCCGTTC	AGGCCGAGAT	GTTCATCCAG	AACGTTATCC	3120
	TGGTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCTT	GTGGCAGTGG	3180
	GGCCCCCTGT	CATCTCTTTT	TCAGTCTCTG	ACATTGTCTC	CAGGGTCTCT	ATTGGGAGGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCACTCAC	CTTTCTCTCT	CCACATCAAG	TCCAGCATAC	3300
25	AGGSCCTTGC	CACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTCCTGCAC	AGATACCAGG	3360
	AGCTGTGAGA	TGACAAACCA	GCTCCTTTTT	TTTTGTTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGGGCTG	GGACCTCATC	AGCATCGCCC	TCATCACCAC	CACGGGGCTG	ATGATCGTTC	3480
	TTATGCAACG	GCAGATTCCC	CCAGCCTATG	CGGGTCTCGC	CATCTCTTAT	GCTGTCCAGT	3540
	TAAACGGGCT	GTTCCAGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTACCTT	3600
30	CGGTGGAGAG	GATCAATCAC	TACATTAAAG	CTCTGTCCCT	GGAAGCACCT	GCCGAATTA	3660
	AGAACAGGCG	TCCCTCCCTT	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGCGA	3720
	AGATGAGGTA	CCGAGAAAAA	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAA	3780
	CTAAAGAGAA	GATTGGCATT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTTCCG	TCTGGTGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAATCA	3900
35	GTGATATTGG	CCTTGGCGAC	CTCCGAAGCA	AACCTCTCTAT	CATTCTCTAA	GAGCGGGTGC	3960
	TGTTCAAGTG	CACGTGTCAG	TCAAATTGGG	ACCCCTTCAA	CCAGTACACT	GAAGACCAGA	4020
	TTTGGGATGC	CTGGAGAGG	ACACACATGA	AAGATGTAT	TGCTCAGCTA	CCTCTGAAC	4080
	TTGAATCTGA	AGTGTGAGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
40	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
	CCATGGACAC	AGAGACAGAC	TTATTGATTC	AAGAGACCAT	CCGAGAAGCA	TTTGACAGCT	4260
	GTACCATGCT	GACCATTGCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTCTCTG	TCCAACGACA	4380
	GTTCCCGATT	CTATGCCATG	TTTGCTGCTG	CAGAGAACA	GGTCCGCTGC	AAGGGCTGAC	4440
45	TCCTCCCTGT	TGACGAAGTC	TCTTTTCTTT	AGAGCAITGC	CATTCCCTGC	CTGGGGCGGG	4500
	CCCTCATGCG	GCTCTCTCTA	CCGAAACCTT	GCCTTTCTCG	ATTTTATCTT	TCGCACAGCA	4560
	GTTCGGGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATTGTATT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTT	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
50	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
	TATTAAATAA	AGCACTGTGC	TAATAACAGT	GCATATTCTT	TTCTATCATT	TTTGACAGT	4860
	TGCTGTACTG	AGAGATCTGG	TTTGTCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCATTCTT	4920
	CTCTAGCTGG	TGGTTTCAAG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCCTCTGCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
55	GAGACGGGTG	GGCGGCTGGA	GACCATGCAG	AGCGCGGTGA	GTTCCTCAGG	CTCCTGCCTT	5100
	CTGTCTGTGT	GTCACTTACT	GTTCCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCCCT	5160
	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTCGCTC	TCTCTCTTTT	GCTGTGTGTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	GCTGGTTCTT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340
60	GTGGGTTCCA	AGCCCTGGAG	CCAACTGCTG	CTTTTGTAGG	TGGCACTTTT	TCATTGTGCTT	5400
	ATTCCACAC	CTCCACAGT	CAGTGGCAGG	GCTCAGGATT	TGCTGGGTCT	GTTCCTCTTT	5460
	CTCACGCGAG	TGCTGCGACA	GTCTCTCTCT	CTCTCTCCCC	TCAAAGTCTG	CAACTTTAAG	5520
	CAGCTCTTGC	TAATCAGTGT	CTCACACTGG	CGTAGAAGTT	TTTGTACTGT	AAAGAGACCT	5580
	ACCTCAGGTT	GCTGTTTGTG	GTGTGGTTTG	GTGTGTTCCC	GCAAACCCCC	TTTGTGCTGT	5640
65	GGGGCTGGTA	GCTCAGGTGG	GCCTGGTCAC	TGCTGTCTAT	AGTTGAATGC	TCAGCGTTGC	5700
	ATGTGCTGAC	CAACTAGACA	TTCTGTGCCC	TTAGCATGTT	TGCTGAACAC	CTTGTGGAAG	5760
	CAAAAATCTG	AAAATGTGAA	TAAAATTATT	TTGATTTTTG	TAAAAAATAA	AAAAAATAA	5820

Seq ID NO: 585 Protein sequence
Protein Accession #: NP_005679.1

	1	11	21	31	41	51	
75	MDIDIGKEY	IIPSPGYRSV	RERTSTSGTH	RDREDSKFR	TRPLECQDAL	ETAARAEGLS	60
	LDASMHSOLR	ILDEEHFKGK	YHGLSALKP	IRTTSKHQHP	VDNAGLFSCM	TPSWLSSLAR	120
	VAHKKGELSM	EDVNSLSKHE	SSDVNCRLE	RLWQELNEV	GPDAASLRV	VWIFCRTRLI	180
	LSIVCLMITQ	LAFSPGPAFM	VKHLEYTQA	TESNLQYSL	LVLGLLLTEI	VRWSLALTW	240
	ALNYRTGVRL	RGAILTMAPK	KILRLKNIKE	KSLGELINIC	SNDGQRMFEA	AAVGSLLAGG	300
80	PVVAIILGMYI	NVILIGTFG	LGSVAVFILFY	PAMMFASRLT	AYFRKRCVAA	TDERVQKME	360
	VLTYIKFIKM	YAWKAPSQS	VQKIREEEERR	ILEKAGYFQG	ITVGVAPIV	VIAVSVPFSV	420
	EMTLGFDLTA	AQAFVTVTVE	NSMTFALKVT	PFSVKSLSEA	SVAVDRPKSL	FLMEEVEMIK	480
	NKPASPHIKI	EMKNATLAWD	SSHSIIQNSP	KLTPKMKDK	RASRGKKEKV	RQLQRTEHQA	540
	VLABQKHLIL	LDSDERPSP	EEEGKHILG	HLRLQRTLHS	IDLEIQBGKL	VGICGSVSGS	600
	KTSLISAILG	QMTLLBGSIA	ISGTFAYVAQ	QAWILNATLR	DNILPGKEYD	EBRYNSVLNS	660
85	CCLRPDLATL	PSSDLTEIGE	RGANLSGGQR	QRISLARALY	SDRSIYILDD	PLSALDAHV	720
	NHIFNSAIRK	HLKSKTVLFW	THQLQYLVD	DEVIFMKEGC	ITERGTHEEL	MNLNGDYATI	780
	FNNLLGETP	FVEINSKKET	SGSQKKSQDK	GPKTGSVKKE	KAVKPEBQQL	VQLEEKQGGS	840

5 VPWSVYGVYI QAAGGFLAPL VIMALPMLNV GSTAPSTWML SYWIKQSGSN TTVTRGNETS 900
VSDSMKDNPH MQYYASIIYAL SMAVMLILKA IRGVVVFVKG LRASSRLHDE LFRRLRSPM 960
KFPDITPTFR ILNRPFKDMD EVDVRLPFQA EMFIQNVILV FFCVGMIAV PFWFLVAVGP 1020
LVILFSVLHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QEPFLHRYQEL 1080
LDDNQAPFFL PTCAMRWLAV RLDLISIALI TTTGLMIVLM HQQIPPAYAG LAISYAVQLT 1140
GLFQFTVRLA SETEARFTSV ERINHYIKTL SLEAPARIKN KAPSPDWQPE GEVTFENAEM 1200
RYRENLPVLV KKVSPTIKPK EKIGIVGRGT SGKSSLGMAL FRLVELSGGC IKIDGVRIED 1260
IGLADLRSLK SIIPQEPVLF SGTVRSNLDP FNQYTEDQIN DALERTHMKC CIAQLPLKLE 1320
10 SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLLOQE TIREAFADCT 1380
MLTIAHRLHT VLGSDRIMVL AQGVVEFDT PSVLLSNDSS RPYAMFAAAE NKVAVKGG

Seq ID NO: 586 DNA sequence
Nucleic Acid Accession #: NM_001327.1
Coding sequence: 89..631

15 1 11 21 31 41 51
AGCAGGGGCG GCTGTGTGTA CCGAGAATAC GAGAATACCT GGTGGGCCCT GACCTTCTCT 60
CTGAGAGCGG GGCAGAGGCT CCGAGAGCAT GCAGGCCGAA GCGCGGGCCA CAGGGGGTTC 120
20 GACGGGCGAT GCTGATGGCC CAGGAGGCCC TGGCATTCTCT GATGGCCAG GGGGCAATGC 180
TGGCGGCCCA GGAAGAGCGG GTGCCACGGG CCGGAGAGGT CCGCGGGGCG CAGGGGCGAG 240
AAGGGCTCCG GGGCGGGGAG GAGGCGCCCC CCGGGGTCCG CATGGCGGCG CGGCTTCAGG 300
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGAG AGCCGCTGCT TTGAGTTCTA 360
25 CCTCGCCATG CCTTTCCGCA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGGCCA 420
GGATGCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACTG TGTCCGCCAA 480
CATACTGACT ATCCGACTGA CTGCTGCAAG CCACCGCCAA CTGCAGCTCT CCATCAGCTC 540
CTGTCTCCAG CAGCTTTCCG TGTTGATGTG GATCAACGAG TGCTTTCTCG CCGTGTTTTT 600
GGCTCAGCCT CCTCAGGGC AGAGGCGCTA AGCCAGCCT GCGGCCCTT CCTAGGTCAT 660
30 GCCTCTCCCG CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720
GTTTGTGCTG GAGGAGGAC GGCTTACATG TTTGTTTCTG TAGAAAATAA AACTGAGCTA

Seq ID NO: 587 Protein sequence
Protein Accession #: NP_001318.1

35 1 11 21 31 41 51
MQAEGRGTTG STGDADGPGG PGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60
PRGFPHGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EAEIARRSLA QDAPFLVPFG 120
40 VLLKEFTVSG NULTIRLTA DHRQLQLSIS SCLQLLSLLM WITQCFLVPF LAQPPSGQRR

Seq ID NO: 588 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 52..459

45 1 11 21 31 41 51
CCTGCTGGGC CCTGACCTTC TCTCTGAGAG CCGGCAGAG GCTCCGAGC CATGCAGGCC 60
GAAGGCGCAG GCACAGGGGG TTGACGGGC GATGCTGATG GCCCAGGAGG CCTTGGCATT 120
50 CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CCGGTGCCAC GGGCGGCAGA 180
GGTCCCGGGG GCGCAGGGGC AGCAAGGGCC TCGGGGCCGA GAGGAGGCGC CCGCGGGGT 240
CCGATGCGGC GTGCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTGCGGGGC CAGGAGGCCG 300
GACAGCCGCG TGCTTCAGTT CCGACTGACT GCTGCAGACC ACCGCCAACT GCAGCTCTCC 360
ATCAGCTCCT GTCTCCAGCA GCTTTCCCTG TTGATGTGGA TCAGCAGTG CTTTCTGCC 420
GTGTTTTTGG CTGAGCTCC CTCAGGGCAG AGGCGCTAAG CCCAGCCTG GCGCCCTTCC 480
55 TAGGTCATGC CTCTCCCTC AGGGAATGGT CCCAGCACGA GTGGCCAGTT CATTGTGGGG 540
GCCTGATTGT TTGTGCTGG AGGAGGACGG CTTACATGTT TGTTCCTGTA GAAAATAAAG 600
CTGAGCTA

Seq ID NO: 589 Protein sequence
Protein Accession #: Eos sequence

60 1 11 21 31 41 51
MQAEGGQGTG STGDADGPGG PGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60
65 PRGFPHGAAS AQDGRCPGCA RRPDSRLLOF RLTAADHRQL QLSISSCLQQ LSLLMWITQC 120
FLPVFLAQAP SQRR

Seq ID NO: 590 DNA sequence
Nucleic Acid Accession #: NM_005562.1
Coding sequence: 90..3671

70 1 11 21 31 41 51
ACAGCGGAGC GCAGAGTGAG AACCAACCAAC CGAGGCGCGG GGCAGCGACC CCTGCAGCGG 60
75 AGACAGAGAC TGAGCGGGCC GGCACCGCCA TGCCCTGCGCT CTGGCTGGGC TGCTGCCTCT 120
GCTTCTCGCT CCTCTGCCC GCAGCCCGGG CCACCTCCAG GAGGGAAGTC TGTGATTGCA 180
ATGGGAAGTC CAGGCAGTGT ATCTTTGATC GGGAACTTCA CAGACAACT GGTAATGGAT 240
TCCGCTGCGT CACTGCAAT GACAACACTG ATGGCATTCA CTGCGAGAAG TGCAAGAATG 300
GCTTTTACCG GCACAGAGAA AGGAGCCGCT GTTTCGCCCT CAATTGTAACT TCCAAAGGTT 360
80 CTCTTAGTGC TCGATGTGAC AACTCTGGAC GGTGCAGCTG TAAACCAAGT GTGACAGGAG 420
CCAGATGCGA CGATGTCTG CCAGGCTTCC ACATGCTCAC GGATGCGGGG TGCACCCAG 480
ACCAGAGACT GCTAGACTCC AAGTGTGACT GTGACCCAGC TGGCATCGCA GGGCCCTGTG 540
ACGCGGGCGC CTGTGTCTGC AAGCCAGCTG TTACTGGAGA ACGCTGTGAT AGGTGTGAT 600
CAGGTTACTA TAACTGTGAT GGGGGGAACC CTGAGGGCTG TACCCAGTGT TCTGTCTATG 660
85 GGCAATCAGC CAGCTGCGCG AGCTCTGAG AATACAGTGT CCATAAGATC ACCTCTACCT 720
TTCATCAAGA TGTGTATGCG TGGAGGCTG TCCAAAGAAA TGGGTCTCCT GCAAAGCTCC 780
AATGGTCACA GCGCCATCAA GATGTGTTTA GCTCAGCCCA ACGACTAGAC CCTGTCTATT 840

	TTGTGGCTCC	TGCCAAATTT	CTTGGGAATC	AACAGGTGAG	CTATGGGCAA	AGCCTGTCTT	900
	TTGACTACCG	TGTGGACAGA	GGAGGCAGAC	ACCCATCTGC	CCATGATGTG	ATTCTGGAAG	960
	GTGCTGGTCT	ACGGATCACA	GCTCCCTTGA	TGCCACTTGG	CAAGACACTG	CCTTGTGGGC	1020
	TCACCAAGAC	TTACACATTG	AGGTTAAATG	AGCATCCAAG	CAATAATTGG	AGCCCCCAGC	1080
5	TGAGTTACTT	TGAGTATCGA	AGGTTACTGC	GGAATCTCAC	AGCCCTCCGC	ATCCGAGCTA	1140
	CATATGGAGA	ATACAGTACT	GGGTACATTG	ACAATGTGAC	CCTGATTTC	GCCCCCCTG	1200
	TCTCTGGAGC	CCAGCACCCT	TGGGTTGAAC	AGTGTATATG	TCCTGTTGGG	TACAAGGGGC	1260
	AAITCTGCCA	GGATTGTGCT	TCTGGCTACA	AGAGAGATTG	AGCGAGACTG	GGGCCTTTTG	1320
	GCACCTGTAT	TCCTTGTAA	TGTCAAGGGG	GAGGGGCTG	TGATCCAGAC	ACAGGAGATT	1380
10	GTTATTTCAG	GGATGAGAAT	CCTGACATTG	AGTGTGCTGA	CTGCCCAATT	GGTTTCTACA	1440
	ACGATCCGCA	CGACCCCGCG	AGCTGCAAGC	CATGTCCCTG	TCATAACGGG	ITCAGCTGCT	1500
	CAGTGATGCC	GGAGACGGAG	GAGGTGGTGT	GCAATAACTG	CCCTCCCGGG	GTACCCGGTG	1560
	CCCGCTGTGA	GCTCTGTGCT	GATGGCTACT	TTGGGGACCC	CTTTGGTGAA	CATGGCCCG	1620
	TGAGGCGCTT	TCAGCCCTGT	CAATGCAACA	ACAATGTGGA	CCCCAGTGCC	TCTGGGAATT	1680
15	GTGACCGGCT	GACAGGCAGG	TGTTTGAAGT	GTATCCACAA	CACAGCCGGC	ATCTACTGCG	1740
	ACCAGTGCAG	AGCAGGCTAC	TTCCGGGACC	CAITGTGCTC	CAACCCAGCA	GACAAGTGTC	1800
	GAGCTTGCAA	CTGTAACTCC	ATGGGCTCAG	AGCCTGTAGG	ATGTCGAAGT	GATGGCACCT	1860
	GTGTTTGCAA	GCCAGGATTT	GGTGGCCCCA	ACTGTGAGCA	TGGAGCATT	AGCTGTCCAG	1920
	CTTGCTATAA	TCAAGTGAAG	ATTGAGATGG	ATCAGTTTAT	GCAGCAGCTT	CAGAGAATGG	1980
20	AGGCCCTGAT	TTCAAAGGCT	CAGGGTGGTG	ATGGAGTAGT	ACCTGATACA	GAGCTGGAAG	2040
	GCAGGATGCA	GCAGGCTGAG	CAGGCCCTTC	AGGACATTCT	GAGAGATGCC	CAGATTTTCAG	2100
	AAGGTGCTAG	CAGATCCCTT	GGTCTCCAGT	TGGCCCAAGT	GAGGAGCCAA	GAGAACAGCT	2160
	ACCAGAGCCG	CCTGGATGAC	CTCAAGATGA	CTGTGGAAAG	AGTTCGGGCT	CTGGGAAGTC	2220
	AGTACCAGAA	CCGAGTTCCG	GATACTCACA	GGCTCATCAC	TCAGATGCAG	CTGAGCCTGG	2280
25	CAGAAAGTGA	AGCTTCTCTG	GGAAACACTA	ACATTCTCTG	CTCAGACCAC	TACGTGGGGC	2340
	CAATGGCTTT	TAAAGTCTG	GCTCAGGAGG	CCACAAGATT	AGCAGAAAGC	CACGTTGAGT	2400
	CAGCCAGTAA	CATGAGACAA	CTGACAGGG	AACTGAGGA	CTATTCCAAA	CAAGCCCTCT	2460
	CACCTGTGCG	CAAGGCCCTG	CATGAAGGAG	TCGGAAGCGG	AAGCGTAGC	CCGGAAGGTC	2520
	CTGTGTGCA	AGGGCTTGTG	GAAAAATTGG	AGAAAAACAA	GTCCCTGGCC	CAGCAGTTGA	2580
30	CAAGGGAGGC	CACAGAGGCG	GAAATTGAAG	CAGATAGGTC	TTATCAGCAC	AGCTTCCGCC	2640
	TCCTGGATTG	AGTGTCTCGG	CTTCAGGAG	TCAGTATCA	GTCTTTTCAG	GTGGAAGAA	2700
	CAAGAGGAT	CAAAACAAA	GCGGATTAC	TCTCAACGCT	GGTAACCAAG	CATATGGATG	2760
	AGTTCAAGCG	TACACAAAAG	AATCTGGGAA	ACTGGAAGA	AGAAGCAGAG	CAGCTCTTAC	2820
	AGAATGAAA	AAGTGGGAGA	GAGAAATCAG	ATCAGCTGCT	TTCCCGTGCC	AATCTTGCTA	2880
35	AAAGCAGAGC	ACAGAGAGCA	CTGAGTATGG	GCAATGCCAC	TTTTATGAA	GTGAGAGCA	2940
	TCCTTAAAAA	CCCTCAGAGG	TTTGACCTGC	AGGTGGACAA	CAGAAAAGCA	GAAAGCTGAG	3000
	AAGCCATGAA	GAGACTCTCC	TACATCAGCC	AGAAGGTTTC	AGATGCCAGT	GACAGACCC	3060
	AGCAGCAGA	AAGAGCCCTG	GGGAGCGCTG	CTGCTGATGC	ACAGAGGCCA	AAGAATGGGG	3120
	CCGGGAGGCG	CCTGGAATTC	TCCAGTGAGA	TGGAACAGGA	GATTGGGAGT	CTGAAGTTGG	3180
40	AAGCCATATG	GACAGCAGAT	GGAGCCTTGG	CCATGGAAAA	GGGACTGGCC	TCTCTGAAGA	3240
	GTGAGATGAG	GGAGTGGGAA	GGAGAGCTGG	AAAGGAAGGA	GCTGGAGTTT	GACACGAATA	3300
	TGGATCCAGT	ACAGATGGTG	ATTACAGAAG	CCCAGAAGGT	TGATACCAGA	GCCAAAGAAC	3360
	CTGGGGTTAC	AATCCAAGAC	ACACTCAACA	CATTAGACGG	CCTCTGTCAT	CTGATGGACC	3420
	AGCCTCTCAG	TGTAGATGAA	GAGGGGCTGG	TCTTACTGGA	GCAGAAAGCT	TCCCGAGCCA	3480
45	AGACCCAGAT	CAACAGCCAA	CTGCGGCCCA	TGATGTGAGA	GCTGGAAGAG	AGGGCAGGTC	3540
	AGCAGAGGGG	CCACCTCCAT	TTGCTGGAGA	CAAGCATAGA	TGGGATTCTG	GCTGATGTGA	3600
	AGAACTTGGA	GAACATTAGT	GACAACTGCG	CCCCAGGCTG	CTACAATACC	CAGGCTCTTG	3660
	AGCAACAGTG	AAGCTGCCAT	AAATATTCTT	CAACTGAGGT	TCTTGGGATA	CAGATCTCAG	3720
	GGCTCGGGAG	CCATGTCTAT	TGAGTGGGTG	GGATGGGGAC	ATTGGAACAT	GTTTAAATGG	3780
50	TATGCTCAGG	TCAACTGACC	TGACCCCAT	CCTGATCCCA	TGGCCAGGTG	GTTGTCTTAT	3840
	TGACCATATC	TCCTTGCTTC	CTGATGCTGG	GCAATGAGGC	AGATAGCACT	GGGTGTGAGA	3900
	ATGATCAAGG	ATCTGAGACC	CAAAGAATAG	ACTGGATGGA	AAGACAAACT	GCACAGGCAG	3960
	ATGTTTGCCT	CATAATAGTG	GTAAGTGGAG	TCCTGGAATT	TGGACAAGTG	CTGTTGGGAT	4020
	ATAGTCAACT	TATCTTTTGA	GTAATGTGAC	TAAAGGAAAA	AACTTGAAGT	TTGCCAGGCG	4080
55	ATGAATTTCT	TCTCAATGTC	AGAACAGAGT	GCAACCCAGT	CACACTGTGG	CCAGTAAAAA	4140
	ACTATTGCTC	CATATTGTCC	TCTGCAAGCT	TCTGTGCTGAT	CAGAGTTTCT	CCTACTTACA	4200
	ACCCAGGGTG	TGAACATGTT	CTCCATTTTC	AAGCTGGAAG	AAGTGAGCAG	TGTTGGAGTG	4260
	AGGACCTGTA	AGGCAGGCCC	ATTGAGAGCT	ATGGTGCTTG	CTGGTGCTTG	CCACCTTCAA	4320
	GTCTCGGACC	TGGGCATGAC	ATCCTTTCTT	TTAATGATGC	CATGGCAACT	TAGAGATTGC	4380
60	ATTTTATTA	TACCAATTTC	TACCAGCAAA	GCAATGTTG	GGAAAGTATT	TACTTTTTCG	4440
	GTTTCAAAAG	GATAGAAAAG	TGTGGCTTGG	GCAATGAAAG	AGGTAATAAT	CTCTAGATTT	4500
	ATTAGTCTTA	ATTCAATCTT	ACTTTTTCGA	CACCAAAAAT	GATGCGCATC	AATGTATTTT	4560
	ATCTTATTTT	CTCAATCTCC	TCTCTCTTTC	CTCCACCCAT	AATAAGAGAA	TGTTCTTACT	4620
	CACACTTCAG	CTGGGTCACA	TCCATCCCTC	CATTATCTCT	TCCATCCATC	TTTCCATCCA	4680
65	TTACCTCCAT	CCATCCTTCC	AACATATATT	TATTGAGTAC	CTACTGTGTG	CCAGGGGCTG	4740
	GTGGGACAGT	GSTGACATAG	TCTCTGCCCT	CATAGAGTTG	ATTGTCTAGT	GAGGAAGACA	4800
	AGCATTTTTA	AAAAATAAAT	TTAAACTTAC	AAACTTTGTT	TGTCACAAGT	GGTGTATTAT	4860
	GCAATAACCG	CTTGGTTTGC	AACCTCTTTG	CTCAACAGAA	CATATGTTCG	AAGACCCCTC	4920
	CATGGGGGCA	CTTGAGTTT	GGCAAGGCTG	ACAGAGCTCT	GGGTTGTGCA	CATTTCCTTG	4980
70	CATTCCAGCT	GTCACTCTGT	GCCTTTCTAC	AACTGATTGC	AACAGACTGT	TGAGTTATGA	5040
	TAAACCAAGT	GGGAATTGCT	GGAGGAACCA	GAGGCATCTC	CACCTTGGCT	GGGAAGACTA	5100
	TGGTGTCTGC	TGTCTTCTGT	ATTTCCTTGG	ATTTCCTGTA	AAGTGTTTTT	AAATAAAGAA	5160
	CAATTGTTAG	ATGCC					

Seq ID NO: 591 Protein sequence
Protein Accession #: NP_005553.1

	1	11	21	31	41	51	
80	MPALWLGCCCL	CPSLLLPAA	ATSREVCDC	NGKSRQCIFD	RELHROTNG	FRCLNCNDNT	60
	DGIHCEKCKN	GFYRHRERDR	CLPCNCSKG	SLSARCDNSG	RCSCPKGVGTG	ARCDRLPGF	120
	HMLTDAGCTQ	DQRLDLSKCD	CDPAGIAGPC	DAGRCVCKPA	VTGERCDRCR	SGYYNLDDGN	180
	PEGCTQCFY	GHSASCRSSA	EYSVHKITST	FHQDVGWKA	VQRNGSPAKL	QWSQRHQDVP	240
85	SSAQRDLDPV	FVAPAKFLN	QQVSYGQSL	FDYVRDGR	HPSAHDVILE	GAGLRITAPL	300
	MPLKTLPCG	LTKYTYPLRN	EHPSNNWSPQ	LSYFEYRELL	RNLTLRLIRA	TYGEYSTGYI	360
	DNVTLISARP	VSGAPAPWVE	QCICPVGYKG	QPCQDCASGY	KRDSARLGPF	GTICPCNCQG	420
	GGACDPDTGD	CYSGDENPDI	ECADCPIGFY	NDPDPFRSCK	PCPCHNGFSC	SVMPETEEYV	480

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

CINCPCGVTV ARCELCAVGD FGDPFGEHGP VRPCQPCQCN NNVDPSASGN CDRLTGRCLK 540
 CINHNTAGIYC DQCKAGYFGD PLAPNPADKC RACNCPNMGs EPVGCSDGT CVCKPGFGGP 600
 NCEHGAFFSCP ACYNQVQIQM DQFMQQLQRM EALISKAQGG DGVVFDTELE GRMQQAEQAL 660
 ODILRDAQLS EGASRSLSGLQ LAKVRSQENS YQSRLDDLKM TVERVRALGS QYQNRVRDTH 720
 RLITQMLSL AESASLGNL NIPASDHVVG PNGFKSLAQE ATRLAESHVE SASNMEQLTR 780
 ETEDYSKQAL SLVRKALHEG VSGSGSGSPDG AVVQGLVEKL ETKSLAQQL TREATQAEIE 840
 ADREYQHSIR LLDSVSRLLQ VSDQSPQVEE AKRIKQKADS LSTLVTRHMD EFKRTQKNLG 900
 NWKEEAQQLL QNGKSGREKS DQLLSRANLA KSRAQEALSM GNATFYEVES ILKNLREFDL 960
 QVNRKAEBE EAMKRLSYIS QKVSDASDKT QQAERALGSA AADAQRKNG AGEALEISSE 1020
 IEQEIGSLNL EANYTADGAL AMEKGLASLK SEMREVEGEL ERKELEFDTN MDAVQMVITE 1080
 AQKVDTRAKN AGVTIQDTLN TLDGLLHMD QPLSVDEEGL VLLEQKLSRA KTQINSQLRP 1140
 MMELEERAR QQRGHLHLE TSIDGILADV KNLENIRDNL PPGCYNTQAL EQQ

Seq ID NO: 592 DNA sequence
 Nucleic Acid Accession #: AF101051.1
 Coding sequence: 221.856

1 11 21 31 41 51
 20 GAGCAACCTC AGCTTCTAGT ATCCAGACTC CAGCGCGCGC COGGCGCGCG ACCCCAACCC 60
 CGACCCAGAG CTTCTCCAGC GCGGCGCGCAG CGAGCAGGCG TCCCGCGCCTT AACTTCCTCC 120
 GCGGGGCGCA GCCACTTTCG GGAGTCCGGG TTGCCACCTT GCAAACTCTC CGCCTCTGC 180
 ACCTGCCACC CCTGAGCCAG CGCGGCGCGC CGAGCGAGTC ATGGCCAACG CGGGGCTGCA 240
 GCTGTTGGGC TTAATCTTCG CCTTCTCTGG ATGGATCGGC GCCATCGTCA GCACTGCCCT 300
 25 GCGCCAGTGG AGGATTTACT CCTATGCCGG CGACAACATC GTGACCGCCC AGGCCATGTA 360
 CGAGGGCGCTG TGGATGTCCT GCGTGTCCGA GAGCAGCGGG CAGATCCAGT GCAAACTCTT 420
 TGACTCCCTG CTGAATCTGA GCAGCACATT GCAAGCAACC CGTGGCTTGA TGGTGGTGG 480
 CATCTCTCTG GAGGTGATAG CAATCTTTGT GGCCACCGTT GGCATGAAGT GTATGAAGTG 540
 CTTGGAAGAC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTCTT 600
 30 TCTTGAGGT CTGGCTATTT TAGTTGCCAC AGCATGGTAT GGCAATAGAA TCGTTCAAGA 660
 ATTCTATGAC CCTATGACCC CAGTCAATGC CAGGTACGAA TTTGTCAGG CTCTCTTCC 720
 TGGCTGGGCT GCTGCTTCTC TCTGCTTCTT GGGAGGTGCC CTACTTTGCT GTTCTGTGCC 780
 CCGAAAAACA ACCTCTTACC CAACACCAAG GCCCTATCCA AAACCTGCAC CTTCCAGCGG 840
 35 GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAAAACA TGTGAAACA AACCAGAAAT 900
 GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GTAATCTGAA 960
 GTATGGTATT ACAAAACAAA CAAACAAACA AAAAACCCAT GTGTTAAAT ACTCAGTGCT 1020
 AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080
 TGTATTACT GCTTCCCAT TGAATATCAT ACTCAATGG GGAAGGGGT GCTCCTTAAA 1140
 40 TATATATAGA TATGATATAT TACATGTTTT TCTATTAAAA ATAGACAGTA AAATCTATT 1200
 CTCAATTAGT TGATACTAGC ATACTTAAAA TATCTTAAAA ATAGTAAAT GTATTAAAT 1260
 CCATATTGAT GAAGATGTTT ATTGATATAT TTTCTTTTTC GTCCCTATAT ACATATGTAA 1320
 CAGTCAAAAT CCAATTACTC TTCTTCAATTA GCTTTGGGTG CCTTTGCCAC AAGACCTAGC 1380
 CTAATTTACC AAGATGAAT TCTTTCAATT CTTCATCGGT GCCCTTTTCA TATACTTATT 1440
 45 TTAATTTTTC CCATAATCTT ATAGCACTTG CATCGTTATT AAGCCCTTAT TTGTTTTGTG 1500
 TTTCAATGCT CTCTATCTCC TGAATCTAAC ACATTTTATA GCCTACATTT TAGTTTCTAA 1560
 AGCCAAGAG AATTATTATC AAATCAGAAC TTTGAGGSCA AATCTTTCTG CATGACCAAA 1620
 GTGATAAAT CCTGTGACCC TCCCCACACA ATCCCTGTAC TCTGACCCAT AGCACTCTTG 1680
 TTTGCTTTGA AATATTGAT CCAATTGAGT AGCTGCATGC TGTTCGCCCA GGTGTGTAA 1740
 50 CACAACCTTA TTGATGAAT TTTTAAAGCTA CTTATTCATA GTTTTATATC CCCCTAAACT 1800
 ACCTTTTGTG TCCCATCTCC TTAATTGTAT TGTTTTCCCA AGTGAATTA TCAATGCTTT 1860
 TATATCTCTG TAATAAGGTG TGGTCTGTTT GTCTGAACAA AGTGTAGAC TTTCTGGAGT 1920
 GATAATCTGG TGACAAATAT TCTCTGTGTA GCTGTAAGCA AGTCACTTAA TCTTTCTACC 1980
 TCTTTTTTCT ATCTGCCAAA TTGAGATAAT GATACCTTAA CAGTTAGAAG AGGTAGTGTG 2040
 55 AATATTAAAT AGTTTATATT ACTCTCATTC TTGAAACATG AACTATGCCT ATGTAGTGTG 2100
 TTTATTGCT CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACTT ACACACGTAC 2160
 CTTCACTGTA TTTACTGCTT TCTCTCTCT ACCAGTCTAT TTCCAAGTAA CAAACCTTAC 2220
 ACACATAACT CCAATGTGTT CAGTGCCTTC CTCTCTCTAC CAGTCTATT CCACGAAACA 2280
 AAACTAAGC ACATACCTTC ATGTGGCTCA GTGCTTCTCT CTCTCTACCA GTCTATTTC 2340
 60 ATTCTTTCAG CTGTGCTGTA CATGTTTGTG CTCTGTTCCA TTTTAAACAC TGCTCTTACT 2400
 TTTCCAGCT GTACAGAATG CTATTCACT TGAGCAAGAT GATGTATGGA AAGGGTGTG 2460
 GCACCTGCTG CTGAGACCTT GGATTTGAGT CTGTTGCTA TCAATCACCG TCTGTGTTG 2520
 AGCAAGGCAT TTGCTGCTG TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAAITC 2580
 CTGATCTTCC CACCTCAGG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640
 65 GTGGTTTTGT AATTGAAAAA GTGCTATACT AAGGGAAAGA ATTGAGGAAT TAACTGCATA 2700
 CGTTTTGGTG TTGCTTTTCA AATGTTTGAA AATAAAAAAA TGTAAAGAAA TGGGTTTCTT 2760
 GCCTTAAACA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAACGGAAT 2820
 AAGATTCTGA GGAAGTCTTA TCTTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880
 ACAGATGTA TGGGAAGAAA TAAAGGCTTA CGTGTGTTGA AATCCACAG CAAGGGAGAT 2940
 70 TTTTGAATCA TAATAACTCA TAAGTGCTA TCTGTTCACT GATGCCCTCA GAGCTCTTG 3000
 TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGGTACT TCATAATAAA 3060
 CTACACAGG AAGTCAAGCC ACGTGTCTT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120
 TGCCTTCAA ACCTGAGAAT ATATGCTTTT GGAAGTTAAA ATTTAAATGG CTTTGGCCAC 3180
 ATACATAGAT CTCTGATG TGTTGAGTGA ATTCATGTG GATATCAGT ACCAAACATT 3240
 75 AAAAAAAT TTTATGCCCC AAAATGACCA ACGAAATGT TACAATAGAA TTTATCCAAT 3300
 TTTGATCTTT TTATATTCTT CTACCACACC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360
 TTATAATGG AATTGTGATA AAGCAATACT CTTTTCAT AAATGTTTT TTAATTTAAA 3420
 AAAAGGAAAA AAAAAAAAAA AAA

Seq ID NO: 593 Protein sequence
 Protein Accession #: AAD16433.1

1 11 21 31 41 51
 85 MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMYEGE WMSCVSQSTG 60
 QIQCKVFDL LNLSSLTQAT RALMVVGILL GVIAIFVATV GMKQMKCLEL DEVQKMRNAV 120
 IGGAIPLLAG LAIIVATAWY GNRIVQEFYD PMTFVNARYE FGQALPTGWA AASLCLLGA 180
 LLCCSCPRKT TSYPTPRPYP KPAESSGKDY V

Seq ID NO: 594 DNA sequence
Nucleic Acid Accession #: NM_006180.1
Coding sequence: 352..2820

5
10
15
20
25
30
35
40
45
50
55
60
65
70

1	11	21	31	41	51	
CCCCATTG	CATCTAACAA	GGAATCTGCG	CCCCAGAGAG	TCCCGGACGC	GCCTGGTGGG	60
TGCCCCGGCG	GCCTGGGCGAT	GCAGCGACG	CGCGCGCGGA	GCTCCGAGCA	GGGTAGGCGC	120
CCCCCTGTAA	AGCGGTTGCG	TATGCCGCGA	CCACTGTGAA	CCCTGCCGCC	TGCCGGAACA	180
CTCTTCGCTC	CGGACCACT	CAGCTCTGTA	TAAGCTGGAC	TGGGACGCCC	CGCAACAAGC	240
ACCGAGGAGT	TAAGAGAGCC	GCAAGCGCAG	GGAAGGCCTC	CCCGCACGGG	TGGGGGAAAG	300
CGGCGGTCG	AGCGCGGGGA	CAGGCACTCG	GGCTGGCACT	GGCTGCTAGG	GATGTCGTCC	360
TGGATAAGGT	GGCATGGACC	CGCCATGGCG	CGGCTCTGGG	GCTTCTGCTG	GCTGGTTGTG	420
GGCTTCTGGA	GGGCGGCTTT	CGGCTGTCCC	ACGTCCTGCA	AATGCAGTGC	CTCTCGGATC	480
TGGTGCAGCG	ACCTTCTCC	TGGCATCGTG	GCAATTCGGA	GATTGGAGCC	TAACAGTGTA	540
GATCCTAGAG	ACATCAACGA	AATTTTCATC	GCAAAACCAGA	AAAGGTTAGA	AATCATCAAC	600
GAAAGATGATG	TTGAAGCTTA	TGTGGGACTG	AGAAATCTGA	CAATTGTGGA	TTCTGGATTA	660
AAATTTGTGG	CTCATAAAGC	ATTTCTGAAA	AACAGCAACC	TGCAGCACAT	CAATTTTACC	720
CGAAACAAAC	TGACGAGTTT	GTCTAGGAAA	CATTTCGGTC	ACCTTGACTT	GTCTGAACCTG	780
ATCCTGTGTG	GCAATCCATT	TACATGTCTC	TGTGACATTA	TGTGGATCAA	GACTCTCCAA	840
GAGGCTAAGT	CACTGCCAGA	CACCTCAGGAT	TTGTACTGCC	TGAATGAAAG	CAGCAAGAAT	900
ATTCCTCTGG	CAAACTGTCA	GATACCCAAT	TGTGGTTTGC	CATCTGCAAA	TCTGGCCGCA	960
CCTAACCTCA	CTGTGGAGGA	AGGAAAGTCT	ATCACATTAT	CCTGTAGTGT	GGCAGGTGAT	1020
CGGTTTCTTA	ATATGTATTG	GGATGTTGGT	AACCTGGTTT	CCAAACATAT	GAATGAAACA	1080
AGGCACACAC	AGGCTCCTTT	AAGGATAACT	AACATTTTCT	CGATGACAG	TGGGAAGCAG	1140
ATCTCTTGTG	TGGCGGAAAG	TCTTGTAGGA	GAAGATCAAG	ATTCTGTCAA	CCTCACTGTG	1200
CATTTTGCAC	CAACTATCAC	ATTTCTCGAA	TCTCCAACCT	CAGACCACCA	CTGTGTCATT	1260
CCATTCACTG	TGAAGGCAAA	CCCCAAACCA	GGGCTTCAGT	GGTTCTATAA	CGGGGCAATA	1320
TTGAATGAGT	CCAAATACAT	CTGTACTAAA	ATACATGTTA	CCAATCACAC	GGAGTACCAC	1380
GGCTGCCTCC	AGCTGGATAA	TCCCACTCAC	ATGAACAATG	GGGACTACAC	TCTAATAGCC	1440
AAGAAATGAGT	ATGGGAAGGA	TGAGAAACAG	ATTTCTGCTC	ACTTCATGGG	CTGGCTTGGG	1500
ATTGACGATG	GTGCAAAACC	AAATTATCCT	GATGTAATTT	ATGAAGATTA	TGGAACGTCA	1560
GCGAATGACA	TGGGGGACAC	CACGAAACAGA	AGTAATGAAA	TCCCTTCCAC	AGACGTCACT	1620
GATAAAACCG	GTCCGGGAACA	TCTCTCGGTC	TATGCTGTGG	TGGTGATTGC	GTCTGTGGTG	1680
GGATTTTGGC	TTTGTGTAAT	GCTGTTTCTG	CTTAAGTTGG	CAAGACATCT	CAAGTTTGGC	1740
ATGAAAGGCC	CAGCTCCGTT	TATCAGCAAT	GATGATGACT	CTGCCAGCCC	ACTCCATCAC	1800
ATCTCCAAATG	GGAGTAACAC	TCCATCTTCT	TGGGAAGGTG	GCCCAGATGC	TGTCATTATT	1860
GGAAATGACCA	AGATCCCTGT	CATTGAAAAT	CCCCAGTACT	TTGGCATCAC	CAACAGTCAG	1920
CTCAAGCCAG	ACACATTTGT	TCAGCACATC	AAGCGACATA	ACATTGTTCT	GAAAAGGGAG	1980
CTAGGCGAAG	GAGCCTTTGG	AAAAGTGTTC	CTAGCTGAAT	GCTATAACCT	CTGTCTCTGAG	2040
CAGGACAAGA	TCTTGTGTGG	AGTGAAGACC	CTGAAGGATG	CCAGTGACAA	TGCACGCAAG	2100
GACTTCCACC	GTGAGGCGCA	GCTCCTGACC	AACCTCCAGC	ATGAGCACAT	CGTCAAGTTC	2160
TATGGCGTCT	CGGTGGAGGG	CGACCCCTCT	ATCATGTGCT	TTGAGTACAT	GAAGCATGGG	2220
GACCTCAACA	AGTTCTCTCA	GGCACACGGC	CCTGATGCGG	TGCTGATGGC	TGAGGGCAAC	2280
CGGCCCAAGG	AACGAGCGCA	GTGCGAGATG	CTGCATATAG	CCCAGCAGAT	CGCCGCGGGC	2340
ATGGTCTACC	TGGGCTCCCA	GCACCTTCTG	CACCGCGATT	TGGCCACCAG	GAATGCGCTG	2400
GTCCGGGAGA	ACTTGTCTGT	GAAAATCGGG	GACTTTGGGA	TGTCGCGGGA	CGTGTACAGC	2460
ACTGACTACT	ACAGGGTCGG	TGGCCACACA	ATGCTGCCCA	TTGCTGTGAT	GCCTCCAGAG	2520
AGCATCATGT	ACAGAAATTT	CACGACGGAA	AGCGACGCTC	GGAGCCTGGG	GGTGTGTGTTG	2580
TGGGAGATTT	TCACTATGAG	CAACAGCCCC	TGGTACCAGC	TGTCAAAACA	TGAGGTGATA	2640
GAGTGTATCA	CTCAGGCGCG	AGTCTCTCAG	CGACCCCGCA	CGTGCCCCCA	GGAGTGTAT	2700
GAGCTGATGC	GGGCTGTGCT	GCAGCGAGAG	CCCCACATGA	GGAAGAACAT	CAAGGGCATC	2760
CATACCTCTC	TTCAAGAACTT	GGCCAAAGCA	TCTCGGCTCT	ACCTGGACAT	TCTAGGCTAG	2820
GGCCCTTTTC	CCGAGACCGA	TCCTTCCCAA	CGTACTCTCT	AGACGGGCTG	AGAGATGAA	2880
CATCTTTTAA	CTGCGCTGCG	AGGCCACCAA	GCTGCTCTCC	TTCACTCTGA	CAGTATTAAAC	2940
ATCAAGAGCT	CGGAGAAGCT	CTCGAGGGAA	CGAGTGTGTA	CTTCTTCTATC	CATAGACACA	3000
GTATTGACTT	CTTTTGGGCA	TTATCTCTTT	CTCTCTTTCC	ATCTCCCTTG	GTTGTTCTTT	3060
TTTCTTTTTC	TAAATTTTCT	TTTTCTTCTT	TTTTTCTGCT	TTCCCTGCTT	CAGATTCTTT	3120
ACCTTCTGTT	TGAAATCAAT	CTGGCTTCTG	CATTACTATT	AACCTCTGAT	AGACAAAGGC	3180
CTTAACAAAC	GTAATTTGTT	ATATCAGCAG	ACACTCCAGT	TTGCCACCA	CAACTAACAA	3240
TGCTTGTGTT	TATTCCTGCC	TTTGTGTGTT	ATGAAAAAAA	GGGAAAAACA	ATATTTCACT	3300
TAAACTTTGT	CACCTCTGCT	GTACAGATAT	CGAGAGTTTC	TATGGATTCA	CTTCTATTTA	3360
TTTATTATTA	TTACTGTTCT	TATTGTTTTC	GGATGGCTTA	AGCCTGTGTA	TAAAAAGAA	3420
AACCTGTGTT	CAATCTGTGA	AGCCTTTATC	TATGGGAGAT	TAAAAACAGA	GAGAAAGAAG	3480
ATTTATTATG	AAACCGAATA	TGGGAGGAAC	AAAGACAACC	ACTGGGATCA	GCTGGTGTCA	3540
GTCCTTACTT	AGGAAATACT	CAGCAACTGT	TAGCTGGGAA	GAATGTATTC	GGCACCTTCC	3600
CTTGAGGACC	TTTCTGAGGA	GTAAAAAGAC	TACTGGCCTC	TGTGCCATGG	ATGATTCTTT	3660
TCCATCACC	AGAAATGATA	GCGTGCAGTA	GAGAGCAAAG	ATGGCTT		

Seq ID NO: 595 Protein sequence
Protein Accession #: NP_006171.1

75
80
85

1	11	21	31	41	51	
MSSWIRWHP	AMARLWGFCH	LVVGFWRAP	ACPTSCCKCSA	SRIWCSDPSP	GIVAPPRLEP	60
NSVDPENITE	IFIANQKRL	IINEDDVVAY	VGLRNLITVD	SGLKPFVAKA	FLKNSNLQHI	120
NFTRNKLTSL	SRKFRHLDL	SELILVGNPF	TCSCDIMWIK	TLQEAQSSPD	TQDLYCLNES	180
SNIPPLANLQ	IENCGLPSAN	LAAPNLVVEE	GKSITLSCSV	AGDPVENMYW	DVGNLVSKHM	240
NETSHTQSL	RITNLSDDDS	GKQISCAVEN	LVGEDQDSVN	LTVEPAPTIT	FLESPTSDEH	300
WCIPPTVKGN	KIPALQWFFN	GAILNESKYI	CTKIHVTNHT	EYHGCLQLDN	PTHMNGDYT	360
LIANKVEYGD	ERQISAHFMG	WPGIDDGANP	NYPDVYIEDY	GTAANDIGDT	TNRSNEIPST	420
DVTDKRGREH	LSVYAVVVIA	SVVGFCLLV	LFLLKLARHS	KFGMKGPASV	ISNDDDSASP	480
LHHSNGSNT	PSSEGGGEDA	VIIQMTKIPV	IENPQYPGIT	NSQLKPDFTV	QHKRHNIVL	540
KRELGEAGAF	KVFLAECYNL	CPEQDKILVA	VKTLKDAASN	ARKDFHREAE	LLTNLQHEHI	600
VKPYGVCEG	DPLIMVFEYM	KHGDLNKFLL	AHGPDVAVLMA	EGNPPTTELQ	SQMLHIAQOI	660
AAGMVYLAQ	HFVHRDLATR	NCLVGENLLV	KIGDFGMSRD	VYSTDYRVG	GHTMLPIRMW	720

PPESIMYRKP TTRSDVWSLG VVLWEIFTYG KQPHYQLSNN EVIECITQGR VLQRPRTCPQ 780
 EVYELMLGCW QREPHMRKNI KGIHTLLQNL AKASPVYLDI LG

Seq ID NO: 596 DNA sequence
 Nucleic Acid Accession #: AF410899
 Coding sequence: 483..2999

1 11 21 31 41 51
 10 GGGAGCAGGA GCCTCGCTGG CTGCTTCGCT CGCGCTCTAC GCGCTCAGTC CCGGCGGTA 60
 GCAGGAGCCT GGACCCAGGC GCGGCGCGCG GCGGTGAGGC GCGGAGCCCC GGCCTCGAGG 120
 TGCATACCGG ACCCCCAATC GCATCTAACA AGGAATCTGC GCCCAAGAGA GTCCCGGACG 180
 CCGCGGCTCG GTGCCCGGCG CGCGGGGCCA TGCAGGAGCG GCGGCGCGCG AGCTCGGAGC 240
 AGCGGTAGCG CCCCCCTGTA AAGCGGTTG CTATGCGGG ACCACTGTGA ACCCTGCGCG 300
 15 CTGCCGGAAC ACTCTTCGCT CCGGACCAGC TCAGCCTCTG ATAAGCTGGA CTCGGCACGC 360
 CCGCAACAAG CACCGAGGAG TTAAGAGAGC CGCAAGCGCA GGGAGGCGCT CCGCGCACGC 420
 GTGGGGGAAA GCGGCGGTTG CAGCGCGGGG ACAGGCACTC GGGCTGGCAC TGGCTGCTAG 480
 GGATGTGCTC CTGGATAAGG TGGCATGGAC CGGCCATGGC GGGCTCTGG GGCCTCTGCT 540
 20 GGCCTGTTGT GGGCTTCGCG AGGCGCGCTT TCGCCTGTCC CAGCTCTGTC AAATGCACTG 600
 CCTCTCGGAT CTGTGTCAGC GACCCCTTCT CTGGCATCGT GGCATTTCCG AGATTGGAGC 660
 CTAACTAGT AGATCTCTAG AACATCACCG AAATTTTCAT CGCAAAACAG AAAAGGTTAG 720
 AAATCATCAA CGAAGATGAT GTTGAAGCTT ATGTGGGACT GAGAAATCTG ACAAATTGTG 780
 ATTCTGGATT AAAATTTGTG GCTCATAAAG CATTTCTGAA AACAGCAAC CTGCAGCACA 840
 25 TCAATTTTAC CCGAAACAAA CTGACGAGTT TGTCTAGGAA ACATTTCCGT CACCTTGACT 900
 TGTCTGAAC CTGCTGCGTG GGCATCCAT TTACATGCTC CTGTGACATT ATGTGGATCA 960
 AGACTCTCCA AGAGGCTAAA TCCAGTCCAG ACATCTCAGG TTTGTACTGC CTGAATGAAA 1020
 GCAGCAAGAA TATTCCCTGT GCAAACTGCG AGATACCCAA TTGTGGTTTG CCATCTGCAA 1080
 ATCTGGCCGC ACCTAACCTC ACTGTGGAGG AAGGAAAGTC TATCACATTA TCCTGTAGTG 1140
 30 TGGCAGGTGA TCGGTTTCCT AATATGTATT GGGATGTTGG TAACCTGGTT TCCAAACATA 1200
 TGAATGAAAC AAGCCACACA CAGGGCTCCT TAAGGATAAC TAACATTTCA TCGGATGACA 1260
 GTGGGAGACA GATCTCTTGT GTGGCGGAAA ATCTTGTAGG AGAAGATCAA GATTCTGTCA 1320
 ACCTCACTGT AATTCTTGA CCAACTATCA CATTTCTGGA ATCTCCAAC TCAGACCACC 1380
 ACTGGTGCAT TCCATTCCT GTGAAAGGCA ACCCCAAACC AGCGCTTCAG TGGTTCTATA 1440
 35 ACGGGGCAAT ATGGAATGAG TCCAAATACA TCTGTACTAA AATACATGTT ACCAATCACA 1500
 CGGAGTACCA GCGCTGCTGC CAGCTGGATA ATCCCACTCA CATGAACAAT GGGGACTACA 1560
 CTCTAATAGC CAAGAGTAGG TATGGGAAGG ATGAGAAACA GATTTCTGCT CACTTCATGG 1620
 GCTGGCCTGG AATTGAAGAT GGTGCAAAACC CAAATTATCC TGATGTAATT TATGAAGATT 1680
 ATGGAATGCG AGCGAATGAC ATCGGGGACA CCACGAAACG AAGTAATGAA ATCCCTTCCA 1740
 CAGACGTCAC TGATAAAACG GGTGCGGAAC ATCTCTCGGT CTATGCTGTG GTGGTGATTG 1800
 40 CGTCTGGAAT GGGATTTTGC CTTTTGTAA TGCTGTTTCT GCTTAAGTTG GCAAGACACT 1860
 CCAAGTTTGG CATGAAGAT TTCTCATGGT TTGGATTGGG GAAAGTAAAA TCAAGACAAG 1920
 GTGTTGGCCG ACCCTCGGTT ATCAGCAATG ATGATGACTC TGCCAGCCCA CTCATCACA 1980
 TCTCCAAATGG GAGTAACACT CCATCTTCTT CGGAAGGTGG CCCAGATGCT GTCAATTATG 2040
 45 GAATGACCAA GATCCCTGTC ATTGAAAATC CCCAGTACTT TGGCATCACC AACAGTCAGC 2100
 TCAAGCCAGA CAGCTTGTGT CAGCACATCA AGCGACATAA CATTGTTCTG AAAAGGGAGC 2160
 TAGGCGAAGG AGCCTTTGGA AAGTGTTCCT TAGCTGAATG CTATAACCTC TGTCTGAGC 2220
 AGGACAAGAT CTTGGTGCCA GTGAAGACCC TGAAGSATGC CAGTGACAAT GCAOAGAGG 2280
 ACTTCCACCG TGAGGCGGAG CTCCTGACCA ACCTCCAGCA TGAGCACATC GTCAAGTTCT 2340
 50 ATGGCGTCTG CGTGGAGGGC GACCCCTTCA TCATGCTCTT TGAGTACATG AAGCATGGGG 2400
 ACCTCAACAA GTTCTCAGG GCACACGGCC CTGATGCCGT GCTGATGGCT GAGGGCAACC 2460
 CGCCACGCGA ACTGACGCG TCGCAGATGC TGCATATAGC CCAGCAGATC GCGCGGGGCA 2520
 TGGTCTACCT GCGCTCCGAG CACTTCGTGC ACCGCGATTT GCGCACCCAG AACTGCTTGG 2580
 TCGGGGAGAA TCTGTGTTG AAAATCGGGG ACTTTGGGAT GTCCCGGGAC GTGTACAGCA 2640
 55 CTGACTACTA CAGGTCGCTG GGCCACACAA TGCTGCCCAT TCGCTGGATG CCTCCAGAGA 2700
 GCATCATGTA CAGGAAATCT ACGACGGAAA GCGACGCTCG GAGCCTGGGG TCGTGTGTTG 2760
 GGGAGATTTT CACCTATGGC AAACAGCCCT GGTACCGACT GTCAACAAT GAGGTGATAG 2820
 AGTGTATCAC TCAGGGCCGA GTCTGCGAGC GACCCCGCAC GTGCCCCAG GAGGTGTATG 2880
 AGCTGATGCT GGGGTGCTGG CAGCGAGAGC CCCACATGAG GAAGAACAAT AAGGGCATCC 2940
 60 ATACCTCTCT TCAGAACTTG GCCAAGGCAT CTCGCGTCTA CTGGACATT CTAGGCTAGG 3000
 GCGCTTTTCC CAGACCGGAT CCTTCCCAAC GTACTCTCTA GACGGGCTGA GAGGATGAAC 3060
 ATCTTTTAA TGCCTGTGGA GGCCACCAAG CTGCTCTCCT TCATCTGAC AGTATTACA 3120
 TCAAGACTC CGAGAAGCTC TCGAGGGAAG CAGTGTGTAC TTCTTCATCC ATAGACACAG 3180
 TATTGACTTC TTTTGGCAT TATCTCTTTC TCTCTTTCCA TCTCCTTGG TTGTCTCTT 3240
 65 TTCTTTTCTT AAATTTCTT TTTCTTCTT TTTTCTGCT TCCCTGCTTC ACGATTCTTA 3300
 CCTTTCTTT TGAATCAAT TGGCTTCTGC ATTACTATTA ACTCTGCATA GACAAAGGCC 3360
 TTAACAAACG TAATTTGTTA TATCAGCAGA CACTCCAGTT TGCCACCAC AACTAACAT 3420
 GCCTTGTGTT ATTCCTGCCT TTGATGTGGA TGAAAAAAG GGAACAACAA TATTTCACTT 3480
 70 AAACCTTTGC ACTTCTGCT TACAGATATC GAGAGTTTCT ATGGAATCAC TTCTATTAT 3540
 TTATTATTAT TACTGTCTT ATTGTTTTG GATGGCTTAA GCGTGTGTAT AAAAAAGAAA 3600
 ACTTGTGTTT AATCTGTGAA GCTTTATCT ATGGGAGATT AAAACAGAG AGAAAGAAGA 3660
 TTTATTATGA ACCGCAATAT GGGAGGAACA AAGCAACCA CTGGGATCAG CTGGTGTACG 3720
 TCCCTACTTA GGAATACTC AGCAACTGTT AGCTGGGAAG AATGTATTGG GCACCTTCCC 3780
 CTGAGGACCT TTCTGAGGAG TAAAAAGACT ACTGGCCTCT GTGCCATGGA TGATTTCTTT 3840
 75 CCCATCACA GAAATGATAG CGTGCACTAG AGAGCAAGA TGGCTTCCGT GAGACACAAG 3900
 ATGGCGCATA GTGTGCTCGG ACACAGTTT GTCTTGTAG GTTGTGATGA TAGCACTGGT 3960
 TTGTTTCTCA AGCGCTATCC ACAGAACCTT TGTCAACTTC AGTTGAAAG AGGTGGATT 4020
 ATGTCCAGAG CTCATTTCCG GGTCAAGTGG GAAAGCC

Seq ID NO: 597 Protein sequence
 Protein Accession #: AAL67965.1

1 11 21 31 41 51
 85 MSSWIRWHP AMARLWGFCH LVVGFWRAP ACPTSCCKSA SRIWCSDFSP GIVAFPRLEP 60
 NSVDPENITE IFIANQKRL EINEDDVEAY VGLRNLITVD SGLKPVAKHA PLKNSNLQHI 120
 NPTRNKLTSL SRKHFRHLDEL SELILVGNPP TCSCDIMWIK TLQEAQSSPD TQDLYCLNES 180
 SRNIPLANLQ IFNCGLPAN LAAPNLTVB GKSITLSCSV AGDPVPNMYW DVGNLVSKHM 240

5
10
15
20
25
30
35
40

NETSHTQGS L RITNISSDDS GKQISCVAEN LVGEDQDSVN LTVHFAPTIT FLESPTSDDH 300
WCIPFTVKGN PKPALQWFYN GAILNESKYI CTKIHVTNHT EYHGCLQLDN PTHMNGDYT 360
LIAKNEYGKD EKQISAHFMG WFGIDDDGAP NYPDVIVEDY GTAANDIGDT TNRSNEIPST 420
DVTDKTGRSH LSVYAVVVIA SVVGFCLLVM LFLKLARHS KFGMKDPSWF GFGVKSRQG 480
VGASVISND DDSASPLHI SNGSNTSPSS EGGPDAVIIG MTKIPVIENP QYFGITNSQL 540
KPDTPVQHIK RHNVILKREL GEGAFGKVFL AECYNLCPEQ DKILVAVKTL KDASDNARKD 600
PHREALLTN LQHEHIVKFY GVCVEGDPLI MVPEYMKHGD LNKFLRAHGP DAVLMABGNP 660
PTELQSQML HIAQQLAAGM VYLASQHFVH RDLATRNCLV GENLLVKIGD FGMSRDVYST 720
DYYRVGGHFM LPIRMPPES IMYRKFTTES DVWSLGVVLW EIFTYGKQPW YQLSNNEVIE 780
CITQGRVLQR PRTPQEVYE LMLGCWQREP HMRKNIKGIH TLLQNLAKAS PVYLDILG

Seq ID NO: 598 DNA sequence
Nucleic Acid Accession #: AB052906
Coding sequence: 74..814

15
20
25
30
35
40

1 11 21 31 41 51
| | | | | |
AAAACCTTGA GGTGATTCAT CTTCCAGGCT CTCCTTCCAT CAAGTCTCTC CTCCCTAGCG 60
CTCTGGGTCC TTAATGGCAG CAGCCGCCGC TACCAAGATC CTTCTGTGCC TCCCGCTTCT 120
GCTCCTGCTG TCCGCTGTGT CCCGGGCTGG GCGAGCCGAC CCTCACTCTC TTTGCTATGA 180
CATCACCGTC ATCCTTAAGT TCAGACCTGG ACCACGGTGG TGTGCGGTTC AAGGCCAGGT 240
GGATGAAAG ACCTTTCTTC ACTATGACTG TGGCAACAAG ACAGTCACAC CTGTGAGTCC 300
CCTGGGAAG AAACCTAAATG TCACAACGGC CTGGAAGACA CAGAACCAG TACTGAGAGA 360
GGTGGTGAG ATACTTACAG AGCAACTGCG TGACATTGAG CTGGAGAATT ACACACCCAA 420
GGAACCCCTC ACCCTGCAGG CCAGGATGTC TTGTGAGCAG AAAGCTGAAG GACACAGCAG 480
TGGATCTTGG CAGTTCAAGT TCGATGGGCA GATCTTCTCT CTCCTTGACT CAGAGAAGAG 540
AATGTGGACA ACGGTTCAAT CTGGAGCCAG AAAGATGAAA GAAAAGTGGG AGAATGACAA 600
GGTGTGGCC ATGTCTCTCC ATTACTTCTC AATGGGAGAC TGTATAGGAT GGCTTGAGGA 660
CTTCTTGATG GGCATGGACA GCACCTGGA GCCAAGTCCA GGAGCACCAC TCGCCATGTC 720
CTCAGGCACA ACCCAACTCA GGGCCACAGC CACCACCTC ATCCTTTGCT GCCTCCTCAT 780
CATCTCCCC TGCTTCATCC TCCCTGGCAT CTGAGGAGAG TCCTTTAGAG TGACAGGTTA 840
AAGCTGATAC CAAAAGGCTC CTGTGAGCAC GGTCTTGATC AAACCTCGCC TTCTGTCTGG 900
CCAGCTGCC ACGACCTAGG GTGTATGTCC AGTGGCCTCC AGCAGATCAT GATGACATCA 960
TGGACCAAT AGGCTATTCA CTGCTTGAT TCCTTTTGCC AACATTTTA CCAGCAGTTA 1020
TACCTAACAT ATTATGCAAT TTTCTCTTGG TGCTACCTGA TGGAAATCTC GCACCTAAAG 1080
TTCTGGCTGA CTAACCAAGA TATATCATTT TCTTTCTTCT CTTTTTGTTC GGAAATCAA 1140
GTACTTCTTT GAATGATGAT CTCTTTCTTG CAAATGATAT TGTGAGTAAA ATAATCAGT 1200
TAGACTTCAG ACCTCTGGGG ATTCTTTCCG TGTCTGAAA GAGAAATTTT AAATTATTTA 1260
ATAAGAAAAA ATTTATATTA ATGATTGTTT CCTTTAGTAA TTTATTGTTT TGTACTGATA 1320
TTTAAATAAA GAGTCTTATT TCCCAAAAAA AAAAAAAAAA AA

Seq ID NO: 599 Protein sequence
Protein Accession #: BAB61048.1

45
50

1 11 21 31 41 51
| | | | | |
MAAAAAATKIL LCLPLLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGQVDEKT 60
FLHYDCGNKT VTFVSPGLKK LNVTTAWKAQ NPVLREVVDI LTELRLDIQL ENYTPKEPLT 120
LQARMSCEQR AEGSSSSSQW FSPDQIPLL FDSEKRMWTT VHPGARKMKE KWENDKVVAM 180
SFHYPSMGDC IGWLEDFLMG MDSTLEPSAG APLAMSSGTT QLRATATTLI LCCLLIILPC 240
FILEPI

Seq ID NO: 600 DNA sequence
Nucleic Acid Accession #: NM_001898.1
Coding sequence: 57..482

55
60
65
70

1 11 21 31 41 51
| | | | | |
GGCTCTCACC CTCTCTCTCT GCAGCTCCAG CTTTGTGCTC TGCCCTCTGAG GAGACCATGG 60
CCCAGTATCT GAGTACCTCT CTGCTCCTGC TGGCCACCTT AGCTGTGGCC CTGGCCTGGA 120
GCCCAAGGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCGAC CTCAATGATG 180
AGTGGGTACA GCGTGCCCTT CACTTCGCCA TCAGCGAGTA TAACAAGGCC ACCAAAGATG 240
ACTACTACAG ACGTCCCGTG CCGGTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGGTGA 300
ATTACTTCTT CGAGGTAGAG GTGGGCCGCA CCATATGTAC CAAGTCCCAG CCCAATTGG 360
ACAACCTGTG CTTCATGAA CAGCCAGAAC TGCAGAAGAA ACAGTTGTGC TCTTTGAGA 420
TCTACGAAGT TCCCTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCTC 480
AGGATCTGT GCCAGGCCAT TCGCACCAGC CACCACCCAC TCCCACCCCT TGTAGTGCTC 540
CCACCCCTGG ACTGGTGGCC CCCACCTGTC GGGAGGCTTC CCCATGTGCC TGGCCCAAGA 600
GACAGACAGA GAAGGCTGCA GGAGTCTTTT GTTGCTCAGC AGGGCGCTCT GCCTCCCTC 660
CTTCTTCTT GCTTCTAATA GCCGTGGTAC ATGGTACACA CCCCCCACC TCTGCAATT 720
AAACAGTAGC ATGCC

Seq ID NO: 601 Protein sequence
Protein Accession #: NP_001889.1

75
80

1 11 21 31 41 51
| | | | | |
MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWVQRALHF AISEYNKATK 60
DDYRRLPLRV LRARQQTGG VNYFFDVEVG RTICTKSQPN LDTCAFHEQP ELQKKQLCSF 120
EIYEVFWENR RSLVKSRCQE S

Seq ID NO: 602 DNA sequence
Nucleic Acid Accession #: NM_003976.2
Coding sequence: 299..961

85

1 11 21 31 41 51
| | | | | |

CTCTGAGCTT CTCTGAGCCT TGTTTGCTCA TCTGGAAAA GGGGATTAAA CCATTACCT 60
CATGGAGTTG TGAAGAATA GCTGCAAAGC ACCTAACACA TAGTAAGGTT CCCAGTGCAG 120
CTACTTCTGC TGGGTTGAGT CTAGCTGTGT AGGCCCTTGT TTCTCACCCT GGAGAAACTG 180
GGGTGGCAGG CCGGTCCCCC ACAAAGATA ACTCATCTCT TAATTTGCAA GCTGCCTCAA 240
5 CAGGAGGGTG GGGGAACAGC TCAACAATGG CTGATGGGCG CTCTGGTGTG TGATAGAGAT 300
GGAACCTGGA CTGAGAGGCC TCTCCACGCT GTCCCACTGC CCCTGGCCTA GGGGGCAGCC 360
TGCCCTGTGG CCCACCTTGG CCGCTCTGGC TCTGCTGAGC AGCGTGCAG AGGCCTCCCT 420
GGGCTCGCGG CCGCGCAGCC CTGCCCCCGG CGAAGGCCCC CCGCTGTCTC TGGCGTCCCC 480
10 GCGCGGCCAC CTGCGGGGGG GACGCACGGC CCGCTGGTGC AGTGAAGAG CCGCGCGGCC 540
GCCGCGCAG CCTTCTCGGC CCGCGCCCCC GCGCCTGCA CCCCCATCTG CTCTTCCCCG 600
CGGGGGCGCG GGGGCGCGGG CTGGGGGGCC GGGCAGCGCG GCTCGGGCAG CGGGGGCGCG 660
GGGCTGCGCG CTGCGCTCGC AGCTGGTGCC GGTGCGCGCG CTGCGCTGGG GCCACCGCTC 720
CGACGAGCTG GTGCGTTTCC GCTTCTGAGC CCGCTCCTCG CCGCGCGCGC GCTCTCCACA 780
15 CGACCTCAGC CTGGCGAGCC TACTGGGCGC CGGGGCGCTG CGACCGCCCC CGGGCTCCCG 840
GCCGCTCAGC CAGCCCTGCT GCCGACCCAC GCGCTACGAA GGGGTCTCCT TCAATGACGT 900
CAACAGCACC TGGAGAACCCT TGGACCGCCT CTCGCGCACC GCCTGCGGCT GCCTGGGCTG 960
AGGGCTGCGT CGAGGCTTTT GCAGACTGGA CCCTTACCGG TGGCTCTTCC TGCTGGGAC 1020
CCTCCGCGAG AGTCCCACTA GCCAGCGGCC TCAGCCAGGG ACGAAGGCTC CAAAGCTGAG 1080
20 AGGCCCTTAC CGGCTGGTGA TGGATATCAT CCCGAACAG GTGAAGGAC AACTGACTAG 1140
CAGCCCCAGA GCGCTCACCT TCGGATCCC AGCCTAAAA ACACAGAGA CCTCAGCTAT 1200
GGAGCCCTTC GGACCACTT CTACAGACT CTGGCACTGG CCGAGGCTCG AACCTGGGAC 1260
CCCTCCTCTG ATGAACACTA CAGTGGCTGA GGCATCAGCC CCGCGCCGAG CCCTGTAGGG 1320
ACAGCATTTG AAGGACACAT ATTGAGTTG CTGTTGTGAA AGTGCTGTG CTGGAAGTGG 1380
25 CCTGTACTCA CTATGGGAG CTGGCCCC

Seq ID NO: 603 Protein sequence

Protein Accession #: NP_003967.1

1 11 21 31 41 51
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60
PAGHLPGGRT ARWCSSGRARR PFPQPSRPAP PFPAPPSALP RGGRAARAGG PGSRARAAGA 120
RGCLRLSQLV PVRLALGLHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180
35 RPSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 604 DNA sequence

Nucleic Acid Accession #: NM_057091.1

Coding sequence: 783..1445

1 11 21 31 41 51
ACTGCGCGCT GAGAGAAGAA TCGGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60
GGACCCCAAA ATCTGCACGT ACCAGCAGTC AGCCGCCCCA CGCAGGGACC GGCTTACCCC 120
TGCCTCCCGG CCTCTACTCA CTCTTCCCG CCTCGGGCCC GGCCTCCAG CTCTCTACTT 180
45 GCGGTGTCTA CAAACTCAAC TCCCGGTTTC CGTGCTCTC CACCGCTGGA GTTCTCTACT 240
CTCCATATCC GAGGGGGCCC TCCAGCAGTC TACCCCTCTC CCAACCTCGG GGGACCTAGC 300
CAAGCTAGGG GGGACTGGAT CCGACGGGTG GAGCAGCCAG GTGAGCCCGG AAAGGTGGGG 360
CGGGGACGGG GCGCTCCAGC CCCCACCCCG GATCTGTGTG ACGCTGGGCG TGGAAATTGA 420
CACCAGCAGG CTGCGGGGCG GGGCAGGAGG CTGCTGAGGG ATGAGTTTGG GCGCGGCCCG 480
50 CAGACAAGGC CCGGGGCTCT CGCCAGCAGC AGTCCCTCG GGGCCAGACC CTGCTGCCA 540
CCCGGCGCTG GAGCCCCACA CCGAGGGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600
TAAAGAGGC ACTGCCAGGT GTACAGTCTT GGCATGCGC TGTTTGAGCT TCGGGGAGA 660
GCCCAGCACT GGTCCCCGGA AAGGTGCCTA GAAGAACAAG GTGCAGGACC CCGTGTCTGC 720
TCAACAGGAG GGTGGGGGAA CAGCTCAACA ATGGCTGATG GGGCTCTCTG GTGTTGATAG 780
55 AGATGGAATC TGGACTTGA GGCCTCTCCA CGCTGTCCCA CTGCCCTTGG CTTAGGCGGC 840
AGCTGTCCCT GTGGCCCACT CTGGCGGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCTC 900
CCCTGGGCTC CCGCGCCCGC AGCCCTGCCC CCGCGAAGG CCCCCGCGCT GTCTGGGCT 960
CCCCCGCGCG CCACCTGCGG GGGGACGCA CCGCCCGCTG GTGCACTGGA AGAGCCCGGC 1020
60 GCGCGCGCGC CGACCTTCTT CCGCCCGCGC CCGCGCGCGC TGCAACCCCA TCTGCTCTTC 1080
CCCGCGGGGG CCGCGCGCGG CCGGCTGGGG GCGCGGGCAG CCGCGCTCGG GCAGCGGGGG 1140
CGCGGGGCTG CCGCTGCGC TCGCAGCTGG TCGCGGTGCG CCGCTCGCGC CTGGGCCACC 1200
GCTCCGACGA GCTGGTGCST TTCCGCTTCT GCAGCGGCTC CTGCGCGCGC GCGCGCTCTC 1260
CACACGACCT CAGCTCTGGC AGCCTACTGG GCGCCGGGGC CCGCGAGCGC CCGCCGGGCT 1320
CCCGCGCGCT CAGCCAGCCC TGCTGCGGAC CCACGCGCTA CCAAGCGGTC TCCTTCATGG 1380
65 ACGTCAACAG CAGCTGGAGA ACCGTGGACC GCCTCTCCGC CACCGCTTGC GGCTGCTTGG 1440
GCTGAGGGCT CGCTCCAGGG CTTTGACAGC TGGACCCCTA CCGGTGGCTC TTCTGCTCTG 1500
GGACCTCTCC GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGACGSAAG GCGCTCAAAGC 1560
TGAGAGGCCCT CTACCGGTGG GTGATGGATA TCATCCCGGA ACAGGTGAAG GGACACTGA 1620
CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCCTA AAAGACACCA GAGACCTCAG 1680
70 CTATGGAGCC CTTCGGACCC ACTTCTACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG 1740
GGACCCCTCC TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCGCGCC CAGGCCCTGT 1800
AGGGACAGCA TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGCC TGTGCTGGAA 1860
CTGGCCTGTA CTACTCATG GGAGCTGGCC CC

Seq ID NO: 605 Protein sequence

Protein Accession #: NP_003967.1

1 11 21 31 41 51
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60
PAGHLPGGRT ARWCSSGRARR PFPQPSRPAP PFPAPPSALP RGGRAARAGG PGSRARAAGA 120
RGCLRLSQLV PVRLALGLHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180
85 RPSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 606 DNA sequence

Nucleic Acid Accession #: NM_057160.1

Coding sequence: 1..714

1 11 21 31 41 51
5 ATGCCCGGCC TGATCTCAGC CCGAGGACAG CCCCTCCTTG AGTCTCTTCC TCCCAAGGCC 60
CACCTGGGTG CCCTCTTTCT CCCTGAGGCT CCACTTGGTC TCTCCGCGCA GCCTGCCCTG 120
TGGCCACACC TGGCCGCTCT GGCTCTGCTG AGCAGGCTCG CAGAGGCTTC CCTGGGCTTC 180
GGCGCCCGCA GCGCTGCCCC CCGCGAAGGC CCCCCTGCTG TCTGGGCTTC CCGCGCGGCT 240
CACCTGCCGG GGGGAGCGAC GGGCGGCTGG TGCACTGGAA GAGCCCGGCG GCGCGCGGCG 300
10 CAGCCTTCTC GGGCCGCGCC CCGCGCGCTG GCACCCCAT CTGCTCTTCC CCGCGGGGCG 360
GGCGCGCGCG GGGCTGGGGG CCGGGGCGAG CGCGCTCGGG CAGCGGGGCG GCGGGGCTGC 420
CGCCTGCGCT CGCAGCTGGT GCGGGTGGCG GCGCTCGGCC TGGGCGACCG CTCGACGAG 480
CTGGTGGCTT TCCGCTTCTG CAGCGGCTCC TGCGCGCGCG CGCGCTCTCC ACACGACCTC 540
AGCCTGGCCA GCGTACTGGG GCGCGGGGCC CTGCGACCGC CCGCGGGCTC CCGCGCGGTC 600
15 AGCCAGCCCT GCTGCCGACC CACGCGCTAC GAAGCGGTCT CCTTCATGGA CGTCAACAGC 660
ACCTGGAGAA CCGTGGACCG CCTCTCGGCC ACCGCGCTGG GCTGCGCTGG CTGAGGGCTC 720
GCTCCAGGGC TTTCAGAGCT GGACCCCTAC CGGTGGCTCT TCTGCGCTGG GACCCCTCCG 780
CAGAGTCCCA CTAGCGACCG GCCTCAGCCA GGGACGAAGG CCTCAAGCTC GAGAGGCCCC 840
TACCGTGGGG TGATGGATAT CATCCCGAA CAGGTGAAGG GACAACTGAC TAGCAGCCCG 900
20 AGAGCCCTCA CCCTGGCGAT CCCAGCCTAA AAGACACAG AGACCTCAGC TATGGAGCCC 960
TTCCGACCCA CTCTCAGAG ACTCTGGCAC TGCGCAGGCC TCGAACCTGG GACCCCTCCT 1020
CTGATGAACA CTACAGTGGC TGAGGCATCA GCGCCCGGCC AGGCCCTGTA GGGACAGCAT 1080
TTGAAGGACA CATATTGAG TTGCTTGGTT GAAAGTGCCT GTGCTGGAAC TGGCCTGTAC 1140
TCACTCATGG GAGCTGGCCC C

Seq ID NO: 607 Protein sequence

Protein Accession #: NP_476501.1

1 11 21 31 41 51
30 MPGLISARGQ PLLEVLPPQA HLGALFLPEA PLGLSAQPAL WPTLAALALL SSVAEASLGS 60
APRSPAPRSG PFPVLASPAQ HLPGGRTARW CSGRARRPPP QPSRPAPPPP APPSALPRGG 120
RAARAGGPGS RARAAAGARG RLRSQLVFVR ALGLGHRSD LVRFRFCSGS CRRARSPHDL 180
35 SLASLLGAGA LRPPPGSRPV SQPCCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG

Seq ID NO: 608 DNA sequence

Nucleic Acid Accession #: NM_057090.1

Coding sequence: 29..715

1 11 21 31 41 51
40 CTGATGGGCG CTCTGGTGT TGATAGAGAT GGAAGTGGGA CTGGAGGCC TCTCAAGCT 60
GTCCCACTGC CCCTGGCCTA GCGGCGAGGC TCCACTTGGT CTCTCCGCGC AGCCTGCCCT 120
45 GTGGCCACCC CTGGCGGCTC TGGCTCTGCT GAGCAGGCTC GCAGAGGCTC CCTGGGCTC 180
CGCGCCCGCG AGCGCTGCTC CCGCGAAGG CCCCCTGCTG GTCTGGGCTC CCGCGCGGCT 240
CCACCTGCCG GGGGAGCGCA CCGCGCGCTG GTGCACTGGA AGAGCCCGGC GCGCGCGGCC 300
GCAGCCTTCT CGGCGCGGCG CCGCGCGGCC TGCACCCCA TCTGCTCTC CCGCGGGGCG 360
CGCGCGCGCG CCGGCTGGGG GCGCGGCGAG CCGCGCTCGG GCAGCGGGGG CCGCGGGCTG 420
CGCGCTGGCG TCGCAGCTGG TGCGGGTGGG CGCGCTCGCG CTGGGCCACC GCTCGACGA 480
50 GCTGGTGGT TTCCGCTTCT GCAGCGGCTC CTGCGCGCGC GCGCGCTCTC CACACGACCT 540
CAGCCTGGCC AGCCTACTGG GCGCGGGGCG CCGCGACCG CCGCGGGGCT CCGCGCGGCT 600
GCAGCAGCCC TGCTGCGGAC CCAAGCGCTA CGAAGCGGTC TCCTTCATGG AGCTCAACAG 660
CACCTGGAGA ACCGTGGACC GCCTCTCGGC CACCGCTGCG GGCTGCTGG GCTGAGGGCT 720
CGCTCCAGGG CTTTGCAGAG TGGACCCCTA CCGGTGGCTC TTCTGCGCTG GGACCCCTCC 780
55 GCAGAGTCCC ACTAGCCAGC GCGCTCAGCC AGGGACGAAG GCCTCAAGG TGAGAGGCCC 840
CTACCGGTGG GTATGGGATA TCATCCCGGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900
CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960
CTTGGGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACTCG GAGCCCTCC 1020
TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCGCGCC CAGGCGCTGT AGGGACAGCA 1080
60 TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGGC TGTGCTGGAA CTGGCTGTA 1140
CTCACTCATG GAGCTGGGCC CC

Seq ID NO: 609 Protein sequence

Protein Accession #: NP_476431.1

1 11 21 31 41 51
65 MELGLGLST LSHCPWPRRQ APLGLSAQPA LNPPLAALAL LSSVAEASLG SAPRSPAPRE 60
GPPPVLASPA GHLPGGRTAR WCSGRARRPP PQPSRPAPPP PAPPALPRG GRAARAGGPG 120
70 SRARAAGARG RLRSQLVFVR RALGLHRSR ELVFRFCSG SCRRARSPHD LSLASLLGAG 180
ALRPPPGSRP VSQPCCRPTR YEAVSFMDVN STWRTVDRLS ATAAGCLG

Seq ID NO: 610 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..1746

1 11 21 31 41 51
80 ATGCCACTGA AGCATTATCT CCTTTTGTG GTGGGCTGCC AAGCCTGGGG TGCAGGTTG 60
GCCTACCATG GCTGCCCTAG CGAGTGATAC TGCTCCAGGG CCTCCAGGT GAGTGACAC 120
GGGCGACGCA TTGTGGGGGT GCCCACCCCT CTGCGCTGGA AGCCATGAG CTGCAGATC 180
CTCAACACGC ACATCACTGA ACTCAATGAG TCCCGGTTCC TCAATATCTC AGCCCTCATC 240
GCCCTGAGGA TTGAGAAGAA TGAGCTGTGG CGCATCAGCG CTGGGGGCTT CGAAACCTG 300
GGCTCGCTGC GCTATCTCAG CCTCGCCAAC AACAGCTGCG AGGTTCTGCC CATCGGCTC 360
85 TTCCAGGGCC TGGACAGCCT TGAGTCTCTC CTTCTGTCCA GTAAACAGCT GTTGAGATC 420
CAGCGCGCCC ACTTCTCCCA GTGCAGCAAC CTCAGGAGC TGCAGTTGCA CCGCAACAC 480
CTGGAATACA TCCCTGACGG AGCCTTGCAC CACTGTGTAG GACTCAGGAA GCTCAATCTG 540

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

GGCAGAATA GCCTCACCAC CATCTCACCC AGGGCTCTTC AGCACCTGGG CAATCTCCAG 600
GTCTCTCCGC TGTATGAGAA CAGGCTCACG GATATCCCCA TGGGCACCTT TGTATGGCTT 660
GTTAACTCTC AGGAACCTGC TCTACAGCAG AACAGATTG GACTGCTCTC CCCTGGTCTC 720
TTCCACAACA ACCACAACCT CCAGAGACTC TACCTGTCCA ACAACCAAT CTCCAGCTG 780
CCACCAGCA TCTTCATGCA GCTGCCCCAG CTCACCGCTC TTACTCTCTT TGGGAATTCC 840
CTGAAGGAGC TCTCTCTGGG GATCTTCCGG CCCATGCCCA ACCTGCCGGA GCTTTGGCTC 900
TATGACAACC ACATCTCTTC TCTACCCGAC AATGTCTTCA GCAACCTCCG CCAGTTGCAG 960
GTCTCTGATT TTAGGCGCAA TCAGATCAGC TTCTCTCCCC CGGGTGCCCT CAACGGGCTA 1020
ACGGAGCTCT GGGAGCTGTC CCTCCACACC AACGCACTGC AGGACCTGGA CGGGAATGTC 1080
TTCCGATGCT TGGCCAACTC GCAGAACATC TCCCTGCAGA ACAATGCCCT CAGACAGCTC 1140
CCAGGGAATA TCTTGCCCAA CGTCAATGGC CTCATGGCCA TCCAGTGCA GAACRACCAG 1200
CTGGAGAACT TGCCCCCTGG CATCTTOGAT CACCTGGGGA AACTGTGTGA GCTGGGCTG 1260
TATGACAATC CCTGGAGGTG TGACTCAGAC ATCCTTCCCG TCCGCAACTG GCTCTGTCTC 1320
AACCCAGCTA GGTTAGGGAC GACACTGTGA CTTGTGTGTT TCAGCCGAGC CAATGTCCGA 1380
GGCCAGTCCC TCAATTATCAT CAATGTCAAC GTTGTCTGTC CAAGCGTCCA TGTCCCTCAG 1440
GTGCTCTAGT ACCCAGAAAC ACCATGGTAC CCAGACACAC CCAGTTACCC TGACACCACA 1500
TCGGTCTCTT CTACCACTGA GCTAACGAGC CCTGTGGAAG ACTACACTGA TCTGACTACC 1560
ATTGAGGTCA CTGATGACCG CAGCGTTTGG GGCATGACCC AGGCCAGAGC CGGGCTGGCC 1620
ATTGCGGCTA TTGTAATTGG CATGTGCGCC CTGGCTGCTC CCCTGGCTGC CTGCGTGGC 1680
TGTGTGCTGT GCAAGAAGAG GAGCCAAGCT GTCTGTATGC AGATGAAGGC ACCCAATCAG 1740
TGTTAAAGAG GCAGGCTGGA GCAGGGCTGG GGAATGATGG GACTGGAGGA CCTGGGAATT 1800
TCATCTTTCT GCCTCCACCC CTGGGTCCAT GGAGCTTTCC CGTGAATGCT CTTTCTGGCC 1860
CTAGATAAAG GTGTGCTCAC CTCTCTCTGA CTTGCTGAT TCTCCGTAG AGAAGCAGGT 1920
CGTGCCGAC CTTCTACAA TCAGGAAGAT AGATCCAACT GGCCATGGCA AAAGCCCTGG 1980
GGATTTCCGA TCTATACCCC TGGGCTTCTC TCGAGAGGGC TCTTCTCCA AATCTCCCC 2040
ACCTGCTCCT CAAGAACAGC CTTCCCTGCG CCCAGGCCCC CTCCGGGCTC CTGTAGACTC 2100
AGTTAGTCCA CAGCTGCTGC ACTTCGTGGG AATAGTTCTC CGCTGAGATA GCGGCTCTCG 2160
CCTAAGTATT ATGTAGTTG ATTTCCTTTC TTTGTTTCT CTTGTTTGTG CTATGGCTTG 2220
ACCCAGCATG TCCCTCAAA TGAAAGTTCT CCCCTTGATT TCTGCTCTCT GAAGCAGGG 2280
TGAGTTCTCT CCTCAAGAA GACTTCAAA CATTTAACTG GTTCTTAAG AGCGTCAAT 2340
CAGCTGTGTT TTGGGATGTC TATGAAAGAG AGAAGGAAA TCATGCCGCT CAGTTCTCTG 2400
AGACAGAAGA GCGGTATCA GTGTCTCACT TGTGATTTT ATCTGGAATA GGAAGAAACA 2460
CCCCAGCACA GCAAGCTCAG CCTTTTAGAG AAGGATATTT CCAAACTGCA AACTTTGCTT 2520
TGAAAGTTT AGCCCTTAA GGAATGAAAT CATGTAGAAT TTTGGACTTC TAAAAACATT 2580
AAAACTAGT TATTAAATAG GATAGAGAA AGAAATCTGG TGCCCTGGGG TCCCTGTGTT 2640
CACCCCTAGA GTTGTGTTTA AAATTTTAA TTGAAGCATG TGAAGTGATC STGCAGAAAA 2700
GTGGGAACAT GATAGTGTAT GCGTTGGTGG ATTTTCACAA ACTGAACATA CTTGTGTAAT 2760
CAGCATCTAG ACCCAGACCC AGAGCATCAC AAATATCCCC CATCTGGGC TTTTCCAGA 2820
GGAGATGGGG GCTTCTGAAG ATGGACTTAC CTGGGACCTG CCCCCATGA GCCAGGACCG 2880
TCCCCCACA GCTAGCCTGT GCAAGGCCCG CGTGCCGAGG GGTGAGGAG AATATGTGGG 2940
TGTGGACAGG ATGGGAGACT GTGGCCTGAA CAGGAGATTT TATTATATCT GGAGACCCCTG 3000
AGAGACCCCT AGACCTGGGG CACCATGGCT GGCCAGGTCA GAAGCATCTT GACTGCAGAG 3060
GTCCGTGACG CCACACCCCT TCCCTGCCA GCAAGTTGTC TGCGGCTCAT CGGAGGCCCC 3120
TCCGCTCGGA GCCTTCTATG GACGTGATAT GCCTGTATCT GTTTTTAATT TTCATTCTTC 3180
ACTTAGGGGA AGTGAATCG CTCAGAGATG AGATCCTTTA ATTGAAAAAG AAGTGTAACG 3240
GAATCTAGTG TCTTTCTAAT GTGTGTAAT TCTCCATCAA CATCACAGTC AGCTGGCAGC 3300
TGAATTCAG AATCTCACT ACAGCAGGCG ACACGGGGGT ACACCGATGG GTACACATGG 3360
GTCTGGGGGC TCCCTGGAGC TCCCTCTGCG TGTGCTCTGG TTAGGAGTTG AGTTGTTTGC 3420
TCCAGGGTTA TTTCTCTCT CGAGTCACAG TCACACGAAT ACCTGCCCTC TCTGGCTTTC 3480
CTGCTATACA CATATTACA TGGCGCTCAA GAAGTTAGGC TCATGGCAAC GTGTGCTTTC 3540
CTCTGGACAA CTGGCCAGT TTACAGTGAA ATGGAGAATT TCAGGTCTCC AGCTCTGCCC 3600
AGGAAGAAGC TCCAGCTGAC TCCACGGGGA TCTGGAATC CACGACCAAT CCGATCGGC 3660
TCTTATTAGC TCCCGCTCC ACAAGACACC TGTGCTTTGG AAATCCACA CCAATCCCGA 3720
TGGGCTCTTA TTAGCTCCCC GCTCCACAAG ACACCTGTGA TCTGGAATC TACCACCAAT 3780
CCGATCGGCG TCTTATTAGC TCCCGCTCC ACAAGACACC TGTGACATCC TCCAGGGCCA 3840
CAGGAGCAGG TGCTGACGAG TTTTCCCTTC CAGTTCTCTG ACAAAAAGTG TCCAGAGGGC 3900
TGTGCTCAA CACTAGTGCA CTTGTAGCT TTTCAACCTC TGTCCAGGG AATCTAGGAG 3960
AGATGAGGCC GGTGAGATC AAGAGATGTC ATCCCCCAG GGTCTCCAAG GCATTTCCAC 4020
ACTATTGGTG GCACCTGGAG GACATGCACC AAGGCTTGCC AGAGCCAAAC GGAAGTGAGC 4080
CCAGAGCAGC GCACATGAGC ATCACCGGCT GATGCTGGCC TGTGCTGCTT GGTGCCAACA 4140
GGGCGATCCC GGCCCGTACC CCTCCAGACA GGAAGCATGG GTTTGCCACC AGACCTGTGC 4200
GGTCTCTCTG TGAAGTGGCT CCAGATGTCT TGTGCTAG GCACAGTGG GCCAGGGCTG 4260
GAGGGAGGTG GGAACCTCA TCATCCGGTG GGCCTGCCA ATCTTAACCC AGAACCCCTA 4320
GGTATTCTCT GCACTAGCCA TGACATTGGA GCACCTTCTC CTCCAGCCAG AGGCTGACCT 4380
GAGGGCCACT GTCTCAGAT GACACCAACC AGGAGCACCC TAGGTGAGGG GTGAGGGCCC 4440
CCTTATGTGA ACCTCTTGCC TCTTCTTTC TCCCATCAGA GTGGTTGGAT GGAGCCATTG 4500
GCCTCTCTTT CTTCAGCGGG CCCTTCAAOC TCTCTGCACC ATGTTGTCTG GCTGAGGAGC 4560
TACTAGAAAA GCTGAGTGGG GTCTCTTTC CAACAGGATG ATGCATTGTC TCAATTCTCA 4620
GGGCTGGAA TGAAGCGGCT GTCCCCAGA AAGCTGGAGT GGGGTACAGA GTTCAGTTTT 4680
OCTCTCTGTT TACAGCTCCT TGACAGTCCC ACGCCCATCT GGAGTGGGAG CTGGGAGTTA 4740
GTGTTGGAGA AGAAACAACA AAAGCCAAAT AGAACCACTA TTTTAAAAA GTGCTTACTG 4800
TGACAGATA CTCCTCAGC ACTGGAAGTG GATTCTCTCT CTAGCCCTCA GCACCCCTGC 4860
GGTAGGAGTG CCGCTCTAC CCAGTTGTGA TGGGATACAG AGGCACITGC TCTTCTGCAT 4920
GGTGTTCAT AGGCTGGGAG TTTTATTTAT CTCCTCAAAC TTTGTACAAG AGCTCATGGC 4980
TTGTCTTGGG CTTCCTCAT TAAACCAAAG GAAATGGAAG CCAATCCCTT GTTGTCTTCC 5040
TTAGTCTTGG TCATCAGAAC CTCCTTGGT ACCATATAGA TCAAAAGCTT TGTAAACACA 5100
GGAATAAATA AACTCTTCCA TCCCTTAAAG AATAGAATAG TTTGTCCCTC TCATGGGAAT 5160
TGGCTGTAT GTATATTGTT CTCTCTCTT AGAATTTAGA GATACAAGAG TTCTACTTAG 5220
AACTTTTAT GGACCAAT TCCACAACCT TTCAGATGCT GATGTAGAGC TATTTGGGAA 5280
GAACCTTCAA ACTCAGGAG TTTGCAGAGA GCAGACAGCT AGAGATAACT CGGGACCCAG 5340
AGTTGTGCGA CAGATGTTAG ATGTATCTTA GCTTTTAGCC ATAAACCACT CAAAGATTCA 5400
GCCCCAGAT CCCACAGTCA GAACTGAATC TGCGTTGTTG GGAAGCCAGC AGTGGCCTTG 5460
GGAAGGAAGC GATGGCTGTG GTTCAGAGAG GGTGGGCTGG CAAGCCACTT CGGGGAGAAA 5520
CTCCTTCCGC CCCAGTTTC TTCTTCTCTT AAGGAGAGAT TGTCTCACC AACCCGCTGC 5580
CTTATGCTG CTTCAAAGC TAGATCATGT TTGCTTGTCT TAGAGAAATTA CTGCAAAATCA 5640
GCCCCAGTGC TTGGGATGTC ATTTACAGAT TTCTAGGCC TCAGGGTTTT GTAGAGTGTG 5700
AGCCCTGGTG GGCAGGGTTG GGGGCTCTGT CTTCTGCTGG ATGCTGCTTG TAATCCATTT 5760

GGTGTACAGA ATCAACAATA AATAATATAC ATGTAT

Seq ID NO: 611 Protein sequence
Protein Accession #: BAB84587.1

5

1	11	21	31	41	51		
MPLKHYLLLL	VGCQAWGAGL	AYHGCPSECT	CSRASQVECT	GARIVAVPTP	LPWNAMSLQI	60	
LNTHTITELNE	SPFLNLSALI	ALRIEKNELS	RITPGAFRNL	GSLRYLSLAN	NKLQVLPIGL	120	
10	FQGLDSLES	LISNNQLLQI	QPAHFSQCSN	LKELQLHGNH	LEYIPDGAFD	HLVGLTKLNL	180
GKNSLTHISP	RVFQHLGNLQ	VLRLYENRLT	DIPMGTFDGL	VNLQELALQQ	NQIGLLSPGL	240	
FINNNHLQRL	YLSNNHISQL	PPSIQMQLPQ	LNRLTLFGNS	LKELSLGIFG	PMPNLRRLWL	300	
YDNHISLSPD	NVFSNLRQLQ	VLILSRNQIS	FISPGAFNGL	TELRELSLHT	NALQDLDDNV	360	
FRMLANLQNI	SLQNNRLRQL	PGNIFANVNG	LMAIQLQNNQ	LENLPLGIFD	HLGKLCERL	420	
15	YDNPWRCDS	ILPLRNWLL	NQPRLGTDV	PVCFSPANVR	GQSLIINVN	VAVPSVHVPE	480
VPSYPETPWY	PDPSPYDDT	SVSSTELTS	PVEDYDLDLT	IQVTDNRSVW	GMTQAQSGLA	540	
IAAIVIGIVA	LACSLAACVG	CCCCKRSQA	VLMQMKAPNE	C			

20

Seq ID NO: 612 DNA sequence
Nucleic Acid Accession #: XM_098151
Coding sequence: 1..447

1	11	21	31	41	51		
ATGATGCATT	TGCTCAATTC	TCAGGGCTGG	AATGAGCCGG	CTGGTCCCCC	AGAAAAGCTGG	60	
AGTGGGGTAC	AGAGTTTCAGT	TTTCCTCTCT	GTTTACAGCT	CCTTGACAGT	CCCAAGCCCA	120	
10	TCTGGAGTGG	GAGCTGGGAG	TCAGTGTGG	AGAAGAAACA	ACAAAAGCCA	ATTAGAACCA	180
CTATTTTAA	AAAGTGCTTA	CTGTGCACAG	ATACTCTTCA	AGCACTGGAC	GTGGATTCTC	240	
TCTCTAGCCC	TCAGCACCCC	TGCGGTAGGA	GTGCCGCCCTC	TACCCACTTG	TGATGGGGTA	300	
30	CAGAGGCCAT	TGCTCTTCTG	CATGGTGTTC	AATAGGCTGG	GAGTTTATT	TATCTCTTCA	360
AACTTGTAC	AAGAGCTCAT	GGCTTGTCTT	GGGCTTTCGT	CATTAAACCA	AAGGAAATGG	420	
AAGCCATTCC	CCTGTTGCTC	TCCTTAG					

35

Seq ID NO: 613 Protein sequence
Protein Accession #: XP_098151

1	11	21	31	41	51		
MMHLNLSQGW	NEPAGPPESW	SGVQSSVFLS	VYSSLTVPRP	SGVGAGSQCW	RRNKSQLEP	60	
40	LFLKSAVCAQ	LFLKHWIIL	SLALSTPAVG	VPPLPTCDGV	QRHLPLCMVF	NRLGVLPFIS	120
NEVQELMACL	GLSSLNQRKW	KFPFCCSP					

45

Seq ID NO: 614 DNA sequence
Nucleic Acid Accession #: NM_002658.1
Coding sequence: 77..1372

1	11	21	31	41	51		
GTCCCGCAG	CGCCGTCGCG	CCCTCCTGCC	GCAGGCCACC	GAGGCGCGCG	CCGTCTAGCG	60	
50	CCCCGACCTC	GCCACCATGA	GAGCCCTGCT	GGCGGCGCTG	CTTCTCTGCG	TCCTGGTGGT	120
GAGCGACTCC	AAAGCGACGA	ATGAACCTCA	TCAAGTTCCA	TGAACTGTGT	ACTGTCTAAA	180	
TGGAGGAACA	TGTGTGTCCA	ACAAGTACTT	CTCCAACATT	CAGTGTGTGA	ACTGCCCAAA	240	
GAAATTCGGA	GGGCAGCACT	GTGAAATAGA	TAAGTCAAAA	ACCTGCTATG	AGGGGAATGG	300	
TCACCTTTAC	CGAGGAAAGG	CCAGCACTGA	CACCATGGGC	CGGCCCTGCC	TGCCCTGGAA	360	
55	CTCTGCCACT	GTCTTTCAGC	AAACGTACCA	TGCCACAGA	TCTGATGCTC	TTCAGCTGGG	420
OCTGGGAAA	CATAATTAAT	CGAGGAACCC	AGACACCGG	AGGCGACCCCT	GGTGCTATGT	480	
GCAGGTGGCG	CTAAGACCGC	TTGTCCAAGA	GTGCATGGTG	CATGACTGCG	CAGATGGAAA	540	
AAAGCCCTCC	TCTCTCCAG	AAGAATTAAA	ATTTCACTGT	GGCCAAAGA	CTCTGAGGCC	600	
60	CCGCTTTAAG	ATTATTGGGG	GAGAATTAC	CACCATGAG	AACAGCCCT	GGTTTGGGCG	660
CATCTACAGG	AGGACACCGG	GGGGCTCTGT	CACCTAAGTG	TGTGGAGGCA	GCCTCATCAG	720	
CCCTGTCTGG	GTGATCAGCG	CCACACACTG	CTTCATTGAT	TACCCAAAGA	AGGAGGACTA	780	
CATCCTCTAC	CTGGTCTGCT	CAAGGCTTAA	CTCCAACAG	CAAGGGGAGA	TGAAGTTTGA	840	
GSTGGAAGAA	CTCATCTCTAC	ACAAGGACTA	CAGCGCTGAC	ACGCTGTCTC	ACCACAACGA	900	
65	CATTGCCCTG	CTGAAGATCC	GTTCACAGGA	GGGCAGGTGT	GCGCAGCCAT	CCGGGACTAT	960
ACAGACCATC	TGCCCTGCCCT	CGATGTATAA	CGATCCCCAG	TTTGGCACAA	GCTGTGAGAT	1020	
CACTGGCTTT	GGAAAGAGAG	ATTCTACCGA	CTATCTCTAT	CCGGAGCAGC	TGAAAATGAC	1080	
TGTTGTGAAG	CTGATTTCCC	ACCGGAGTGT	TCAGCAGCCC	CACCTACTAG	GCTCTGAAGT	1140	
70	CACCAACAAA	ATGCTATGTG	CTGCTGACCC	CCAATGGAAA	ACAGATTTCCT	GCCAGGGAGA	1200
CTCAGGGGGA	CCCTCTGCTT	GTTCCTTCCA	AGGCGCGATG	ACTTTGACTG	GAATTGTGAG	1260	
CTGGGGCCGT	GGATGTGCCC	TGAAGGACAA	GCCAGGCGTC	TACACGAGAG	TCTCACACTT	1320	
CTTACCCCTG	ATCCGCACTG	ACACCAAGGA	AGAGATGGC	CTGGCCCTCT	GAGGGTCCCC	1380	
AGGGAGGAAA	CGGGCACCAC	CGCTTTTCTT	GCTGGTTGTC	ATTTTTCAG	TAGAGTCATC	1440	
75	TCCATCAGCT	GTAAGAAGAG	ACTGGGAAGA	TAGGCTCTGC	ACAGATGGAT	TTGCCCTGTG	1500
CACCAACAGG	GTGAACGACA	ATAGCTTTAC	CCTCACTGAT	AGGCTGGGT	GCTGGCTGCC	1560	
CAGACCCCTC	GGCCAGGATG	GAGGGGTGGT	CCTGACTCAA	CATGTTACTG	ACCAGCAACT	1620	
TGCTTTTTC	TGGACTGAAG	CCTGCAGGAG	TTAAAAAGGG	CAGGGCATCT	CCTGTGCATG	1680	
GGCTCGAAGG	GAGAGCCAGC	TCCCCCGACC	GGTGGGCATT	TGTGAGGCC	ATGGTTGAGA	1740	
80	AATGAATAAT	TGGCAATTA	GGAAAGTTAA	GCAGCTGAGG	TCTCTTGAGG	GAGCTTAGCC	1800
AATGTGGGAG	CAGCGGTTTG	GGGAGCAGAG	ACACTAACGA	CTTCAGGGCA	GGGCTCTGAT	1860	
ATTCTCATGA	TGTATCGAGG	AATATATATG	TGTGTGTATG	TTTGACACT	TGTTGTGTGG	1920	
GCTGTGAGTG	TAAGTGTGAG	TAAGAGCTGG	TGCTGATTTG	TTAAGTCTAA	ATATTTCTCT	1980	
AAACTGTGTG	GACTGTGATG	CCACACAGAG	TGGTCTTTCT	GGAGAGGTTA	TAGGTCATCT	2040	
CTGGGGCTCT	TTGGGTCCCC	CACGTGACAG	TGCCCTGGAA	TGTAATTTAT	CTGCAGCATG	2100	
85	ACCTGTGACC	AGCACTGTCT	CAGTTTCACT	TTCACTAGA	TGTCCTTTCT	TTGGCCAGTT	2160
ATCCCTTCTT	TTTACCTCTG	TTTATCCAT	CCTCACTGGG	TGGGTGAGG	ACCACTCTCT	2220	
ACACTGAATA	TTTATATTTT	ACTATTTTAA	TTTATATTTT	TGTAATTTTA	AATAAAAGTG	2280	
ATCAATAAAA	TGTGATTTTT	CTGA					

Seq ID NO: 615 Protein sequence
Protein Accession #: NP_002649.1

```

5      1      11      21      31      41      51
|      |      |      |      |      |
MRALLARLLL CULVVSDESKG SNEHQVPSN CDCLNGGTCV SNKYFSNIHW CNCPKKFGGQ 60
HCEIDKSKTC YEGNGHFYRG KASTDTMGRP CLPWN SATVL QQTYAHRRSD ALQLGLGKHN 120
YCRNPDNRRT PCYVQVGLK PLVQECMVHD CADGKKPSSP PEELKFPQCG KTLRPRFKII 180
10    GGEFTTIENQ PWFAYIRRH RGGSVTYVCG GSLISPCWVI SATHCFIDYP KKEDYIVYLG 240
RSRLNSNTQG EMKFEVENLI LHKDYSADTL AHNDIALLK IRSKEGRCAQ PSRTIQTICL 300
PSMYNDPQFG TSCEITGFGK ENSTDVLYPE QLKMTVVKLI SHRECOQPHY YGSEVTTKML 360
CAADPQWKTD SCQDGGSGGFL VCSLQGRMTL TGI VSWGRGC ALKDKPGVYT RVSHFLPWIR 420
SHTKEENGLA L

```

Seq ID NO: 616 DNA sequence
Nucleic Acid Accession #: NM_024422.1
Coding sequence: 202..2907

```

20      1      11      21      31      41      51
|      |      |      |      |      |
CGCCAAAGGA AAAGCCCTTC GGATGAGAGG CAGGCGCTTC AGAGAAGCTA AGAAAAGCAC 60
CTCTCCGCGC GCGCCACCTC CTCGCGCTCG CGCTCCTCCT GAGCAGCGGG CCCAGACTGC 120
GCTCCGCGCG CCGCCCTCGC CCGCGGAGC CCTCCTACCC CGGCCCGACG CTCGCCCGCG 180
25    GACCTGCCCC GAGCCCTCTC CATGAGGCA GCGCGCCCTC CCGGCTCCTG GAACGGAGCC 240
CTCTCGCGCG TGCTCTCTGT GACCTCGCG ATCTTAATAT TTGCCAGTGA TGCCCTGCAA 300
AATGTGACAT TACATGTTCC TCCTCAACTA GATGCCGAGA AACTTGTGG TAGAGTTAAC 360
CTGAAAGAGT GCTTTACAGC TGCAAACTA ATTCATTCAA GTGATCTGTA CTTCCAAATT 420
TTGGAGGATG GTTCAGTCTA TACAACAAAT ACTATTCTAT TGTCTCTGGA GAAGAGAAGT 480
30    TTATCCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAATATT TGTCTTTTGG 540
GAGCATCAAA CAAAGGTCTC AAAGAAAAGA CATACTAAGG AAAAAGTTCT AAGGCGCGCC 600
AAGAGAGAGT GGGCTCCAAAT TCCTTGTTCG ATGCTAGAAA ACTCCTTGGG TCCTTTTCCA 660
CTTTTCTCTC AACAGGTTCA ATCTGACACG GCCCAAACT ATACCATATA CTATTCCATA 720
AGAGGTCTCT GAGTGTACCA AGAACCTCGG AATTTATTTT ATGTGGAGAG AGACACTGGA 780
35    AACTGTGATT GTACTCGTCC TGTAGATCGT GAGCAGTATG AATCTTTTGA GATAATTGCC 840
TTTGCAACAA CTCACAGTGG GTATATCTCA GAACCTCCAC TGCCCTTAAT AATCAAAATA 900
GAGGATGAAA ATGATAACTA CCCAATTTT ACAGAAGAAA CTTATACTTT TACATTTT 960
GAAAATTGCA GAGTGGGCAC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT 1020
GACAGAGTGC ACACAGCGCT GAAGTACTCC ATCATTGGGC AGGTGCCACC ATCACCACCC 1080
40    CTATTTCTTA TGCTTTCCAA TACAGCGCTG ATCACCACAA CATCATCTCA GCTAGACAGA 1140
GAGTTAATTG ACAAGTACCA GTTGAAATA AAAGTACAAG ACATGGATGG TCAGTATTTT 1200
GGTCTACAGA CAACCTCAAC TTGTATCATT AACATTGATG ATGTAAATGA CCACTTGCCA 1260
ACATTTACTC GTACTTCTTA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGAAATC 1320
TTACAGATTA CTGTTGAGGA TAAGGACTTA GTGAATAGT CTAAGTGGAG AGCTAATTAT 1380
45    ACCATTTTAA AGGCGAATGA AATGGCAAT TTTAAATG TAACAGATGC CAAACCAAT 1440
GAAGGAGTTT TTTGTGATG TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTTG 1500
CAAAATGGTG TAGTTAATGA AGCTCCATT TCCAGAGAGG CTAGTCAAG ATCAGCCATG 1560
AGCACAGCAA CAGTTACTGT TAATGTAGAA GATCAGGATG AGGCGCCTGA GTGTAAACCT 1620
50    CCAATACAGA CTGTTGCGAT GAAAGAAAAT GCAGAAAGTG GAACAACAAG CAATGGATAT 1680
AAAGCATATG ACCCGAATGA AAGAAAGTAG AGTGGCATAA GGTATAAGAA ATTAAGTAT 1740
CCAACAGGAT GGGTACCAAT TGATGAAAT ACAGGATCAA TCAAGTTTT CAGAGCCTG 1800
GATAGAGAGG CAGAACCAAT CAAAATGGC ATATATAATA TTACAGTCTT TGCATCAGAC 1860
CAAGGAGGGA GAACATGTAC GGGGACACTG GGCATTATAC TTCAAGACGT GAATGATAAC 1920
AGCCCATCTA TACCTAAAAA GACAGTGATC ATCTGCAAAAC CCACCATGTC ATCTGCGGAG 1980
55    ATTGTTGCGG TTGATCTCTA TGAGCCTATC CATGGCCCAT CCTTTGACTT TAGTCTGGAG 2040
AGTTCTACTT CAGAAGTACA GAGAATGTGG AGACTGAAAG CAATTAATGA TACAGCAGCA 2100
CGTCTTCTCT ATCAGAATGA TCCTCCATT GGCTCATATG TAGTACCTAT AACAGTGAGA 2160
GATAGACTTG GCATGTCTAG TGTCACTTCA TTGGATGTGA CACTGTGTGA CTGCATTACC 2220
GAAAATGACT GCACATCATG TGTAGATCCA AGGATTGGCG GTGGAGGAGT ACAACTTGGA 2280
60    AAGTGGGCAAT TCCTTGCAAT ATTGTTGGGC ATAGCATTGC TCTTTTGCAAT CCTGTTTACG 2340
CTGGTCTGTG GGGCTTCTCG GACGTCTAAA CAACCAAAAG TAATTCCTGA TGATTTAGCC 2400
CAGCAGAACC TAATTGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGCG 2460
AATGGCTTCA CAACCAAAAC TGTGGGCGCT TCTGCTCAGG GAGTTTGTGG CACCGTGGGA 2520
TCAGGAATCA AAAACGGAGG TCAGGAGACC ATCGAAATGG TGAAAGGAGG ACACAGGACC 2580
65    TCGGAATCCT GCGCGGGGGC TGGCCACCAT CACACCTGG ACTCCTGCAG GGGAGGACAC 2640
ACGGAGGTGG ACRAGTGCAG ATACACTTAC TCGGAGTGGC ACAGTTTAC TCAGCCCGCT 2700
CTTGGTGAAA AAGTGTATCT GTGTAATCAA GATGAAAATC ACAAGCATGC CCAAGACTAT 2760
GTCTGACAT ATAACATGA AGGAAGAGGA TGCGTGGCTG GGTCTGTAGG TGTGTGCACT 2820
70    GAACGACAAG AAGAAGATGG GCTTGAATTT TTGGATAATT TGAGGCCCAA ATTTAGGACA 2880
CTAGCAGAAG CATGCATGAA GAGATGAGTG TGTCTAATA AGTCTCTGAA AGCCAGTGGC 2940
TTTATGACTT TTAATAAAAA TTACAAACCA AGAATTTTTT AAAGCAGAAG ATGCTATTTG 3000
TGGGGGTTTT TCTCTCATTA TTTGGATGGA ATCTCTTTGG TCAATGCAC ATTTACAGAG 3060
AGACACTATA AACAGTACA CAAATTTTTC AATTTTACA TATTTTAAA TTACTTATCT 3120
75    TCTATCCAAG GAGGTCTACA GAGAAATTA AGTCTGCCCT ATTTGTACA TTTGGGTATA 3180
ATGACAACAG CCAATTTATA GTGCAATAAA ATGTAATTA TTCAAGTCTT TATTATAGAC 3240
TATTTGAAGC ACAACCTAAT GGAATTTGT AGAGACCTTG CTTTAACATT ATCTCCAGTT 3300
AATTAAGTGT TCATGTGGTG CTTGGAACCT GTTGTTTTCC TGAACATCTA AAGTGTGTAG 3360
ACTGCATTCT TGCTATTATT TTATTCTTGT ATGTGACCTT TTTCACTGTG CAAAGGGAGA 3420
80    TTTCTAGCCA GGCATTGACT ATTACAATTT CATT

```

Seq ID NO: 617 Protein sequence
Protein Accession #: NP_077740.1

```

85      1      11      21      31      41      51
|      |      |      |      |      |
MEARPPSGSW NGALCRLLLL TLAILIFASD ACKNVTLHVP SKLDAEKLVG RVNLKECFPTA 60
ANLIHSSDPD PQLEDSGVY TINTILLSSE KRSFTILLSN TENQEKKIF VFLEHQTKVL 120

```


KKRHTKEKVL RRAKRRWAPI PCSMLENSLG PFPLFLQQVQ SDTAQNTYII YSIRGPGVDQ 180
EPRNLFYVER DTGNLYCTRP VDREQYESFE IIAFATTDPG YTPPLPLPLI IKIEDENDNY 240
PIFTEETVTF TIFENCROVT TVGQVCATDK DEPDTHMTRL KYSIIGQVPP SPTLFSMHPT 300
5 TGVIITTSQ LDRELIDKQV LKIKVQMDG QYFGLQTTST CIINIDVDND HLPTFTTSY 360
VTSVEENTVD VEILVETVED KDLVNTANWR ANYTILKONE NGNFKIVTDA KTNBGLCVV 420
KPLNYEEKQO MILQIGVUNE APPSREASPR SAMSTATVTV NVEDQDEGPE CNPPIQTVRM 480
KENAEVGTTS NGYKAYDPET RSSSGIRYKK LTDPTGWVTI DENTGSIKVF RSLDREAETI 540
KNGIYNITVL ASLQGGRTCT GTLGIILQDV NDNSFFIPKK TVIICKPTMS SAEIVAVDPD 600
10 EPIHGPPFDF SLSSSTSEVQ RMRLLKAIND TAARLSYQND PPGSVYVPI TVRDLGMSS 660
VTSLDVTLCD CITENDCTHR VDPRIIGGGV QLKKNAILAI LLGIALLFCE LFTLVCGASG 720
TSKQPKVIPD DLAQQLNLS NTEAPGDDKV YSANGFTTQT VGASAQGVCG TVSGGIKNGG 780
QETIEMVKGQ HQTSESCRGA GHHTLDSR GRGTEVDNCR YTYSEWHSPT QPRLGEKVYL 840
CNQDENHKGH QDVVLTNYE GRGSVAGSVG CCSERQEEDG LEFLDNLEPK FRTLAEACMK 900
R

Seq ID NO: 618 DNA sequence
Nucleic Acid Accession #: NM_004949.1
Coding sequence: 202..2745

20 1 11 21 31 41 51
CGCCAAAGGA AAAGCCCTT GGATGAGAGG CAGGCGCTTC AGAGAAGCTA AGAAAAAGCAC 60
CTCTCGCGCG GCCCACCTC CTCGCGCTCG CGCTCCTCCT GAGCAGCGGG CCCAGACTGC 120
25 GCTCCGCGCG CGGCCCTCGC CCCGCGGAGC CCTCCTACCC CGGCCCGAGC CTCGCGCGCG 180
GACCTGCCCG GAGCCCTCTC CATGGAGGCA GCCCGCCCTC CGGCTCCTG GAAAGGAGCC 240
CTCTCGCGCG TGCTCTGCTC GACCTCGCG ATCTTAATAT TTGCCAGTGA TGCTGCAAA 300
AATGTGACAT TACATGTTCC CTCCAACTA GATGCCGAGA AACTTGTGG TAGAGTTAAC 360
CTGAAAGAGT GCTTTACAGC TGCAAACTA ATTCATTCAA GTGATCCTGA CTCCAAATT 420
30 TTGGAGGATG GTTCAGTCTA TACAACAAAT ACTATTCTAT TGTCTCTGGA GAAGAGAAGT 480
TTTACCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAAATAT TGTCTTTTG 540
GAGCATCAAA CAAGAGTCTC AAAGAAAAGA CATACTAAAG AAAAAGTTCT AAGGCGGCGC 600
AAGAGAAGAT GGGCTCCAAT TCCTTGTTCG ATGCTAGAAA ACTCCTTGGG TCCTTTTCCA 660
CTTTCTCTTC AACAGGTTCA ATCTGACACG GCCCAAACT ATACCATATA CTATTCCATA 720
35 AGAGGTCCTG GAGTGTACCA AGAACCTCGG AATTTATTTT ATGTGGAGAG AGACACTGGA 780
AATCTGTATT GTACTCGTCT TGTAGATCGT GAGCAGTATG AATCTTTTGA GATAATTGCC 840
TTTGCAACAA CTCACAGTGG GTATACTCCA GAACCTCCAC TGCCCTTAAT AATCAAAATA 900
GAGGATGAAA ATGATTAATA CCCAATTTT ACAGAAGAAA CTTATACTTT TACAATTTT 960
GAAAAATGCA GAGTGGGAC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT 1020
40 GACACGATGC ACACAGCCTC GAAGTACTCC ATCATTGGGC AGGTGCCACC ATCACTCCAC 1080
CTATTTCTTA TGATCTCAAC TACAGGCGTG ATCACCACAA CATCATCTCA GCTAGACAGA 1140
GAGTTAATTG ACAGTACCA GTTGAAAATA AAAGTACAAG ACATGGATGG TCAGTATTTT 1200
GGTCTACAGA CACTCTCAAC TTGTATCATT AACATTGATG ATGTAATAGA CCACCTGCCA 1260
ACATTTACTC GTACTCTTA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGAAATC 1320
TTACGAGTTA CTGTTGAGGA TAAGGACTTA GTGAATACTG CTAAGTGGAG AGCTAATTAT 1380
45 ACCATTTTAA AGGCAATGA AAATGGCAAT TTTAAATGG TAACAGATGC CAAACCAAT 1440
GAAGGAGTTC TTTGTGTAGT TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTTG 1500
CAAAATGGTG TAGTTAATGA AGCTCCATT TCCAGAGAGG CTAGTCCAAG ATCAGCCATG 1560
AGCACAGCAA CAGTTACTGT TAATGTAGAA GATCAGGATG AGGGCCCTGA GTGTAACCTT 1620
50 CCAATACAGA CTGTTGCGAT GAAAGAAAAT GCAGAAAGTG GAAACAACAG CAATGGATAT 1680
AAAGCATATG ACCCAGAAAC AAGAAGTAGC AGTGGCATAA GGTATAAGAA ATTAAGTAT 1740
CCACAGAGGT GGGTCAACAT TGATGAAAAT ACAGGATCAA TCAAGTTT CAGAAGCCTG 1800
GATAGAGAGG CAGCAACCAT CAAAAATGGC ATATATAATA TTACAGTCTC TGATCAGAC 1860
CAAGGAGGGA GAACATGTAC GGGGACACTG GGCATTATAC TTCAAGACGT GAATGATAAC 1920
AGCCCATTTA TACTTAAAAA GACAGTGATC ATCTGCAAAAC CCACCATGTC ATCTGCGGAG 1980
55 ATTGTGCGG TTGATCTCTA TGAGCCTATC CATGGCCCAAC CCTTTGACTT TAGTCTGGAG 2040
AGTTCTACTT CAGAAGTACA GAGAATGTGG AGACTGAAAG CAATTAATGA TACAGCAGCA 2100
CGTCTTTCTC ATCAGAAATG TCCTCCATT TGCCTCATAT TAGTACCTAT AACAGTGAGA 2160
GATAGACTTG GCATGCTAG TGTCACTTCA TTGGATGTTA CACTGTGTGA CTGCAATTACC 2220
GAAAATGACT GCACACATCG TGTAGATCCA AGGATTGGCG GTGGAGGAGT ACAACTTGA 2280
60 AAGTGGGCA TCCTTCAAT ATTGTGGGC ATAGCATTGC TCTTTTGCAT CCGTTTACG 2340
CTGGTCTGTG GGGCTTCTCG GACGTCTAAA CAACCAAAAG TAATTCTGA TGAATTAGCC 2400
CAGCAGAAC TAATTGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGG 2460
AATGGCTTCA CAACCAAAAC TGTGGGCGCT TCTGCTCAGG GAGTTTGTGG CACCGTGGGA 2520
TCAGGAATCA AAAACGAGG TCAGGAGACC ATCGAAATGG TGAAGAGAGG ACACAGACC 2580
65 TCGGAATCCT GCGGGGGGGG TGGCCACCAT CACACCTGAG ACTCCTGCGG GGGAGGACAC 2640
ACGGAGGTGG ACAAATGCG ATACACTTAC TCGGAGTGGC ACAGTTTAC TCAGCCCGT 2700
CTTGGTGAAG AATCCATTAG AGGACACACT CTGATTAAAA ATTAACAAT GAAAGAAAGT 2760
GTATCTGTGT AATCAAGATG AAAATCACAA GCATGCCCAA GACTATGTCC TGACATATAA 2820
70 CTATGAAGGA AGAGGATCGG TGGCTGGGTC TGTAGGTTGT TGCAGTGAAC GACAAGAAGA 2880
AGATGGGCTT GAATTTTGG ATAATTTGGA GCCCAATTT AGGACACTAG CAGAAGCATG 2940
CATGAAGAGA TGAGTGTGTT CTAATAAGTC TCTGAAAGCC AGTGGCTTTA TGACTTTTAA 3000
AAAAAATTAC AAACCAAGAA TTTTAAAG CAGAAGATGC TATTTGTGG GGTTTTCTC 3060
TCATTATTGG GATGGAATCT CTTTGGTCAA ATGCACATT ACAGAGAGAC ACTATAAACA 3120
75 AGTACACAAA TTTTCAATT TTTACATATT TTTAAATTAC TTATCTCTA TCCAAGGAGG 3180
TCTACAGAGA AATTAAGTC TGCCTTATTT GTTACATTG GGTATAATGA CAACAGCAA 3240
TTTATAGTGC AATAAATGT AATTAATTCA AGTCTTATT ATAGACTATT TGAAGCACA 3300
CCTAATGGAA AATTGTAGAG ACCTTGCTTT AACATTATCT CCAGTTAAT AAGTGTTCAT 3360
GTGGTCTGTG GAACTGTGT TTTCTGAA CATCTAAGG GTGTAGACTG CATTCTGTCT 3420
80 ATTATTTTAT TCTGTAAATG TGACCTTTTC ACTGTGCAA GGGAGATTTC TAGCCAGGCA 3480
TTGACTATTA CAATTCATT

Seq ID NO: 619 Protein sequence
Protein Accession #: NP_004940.1

85 1 11 21 31 41 51
MEARPSGSW NGALCRLLLL TLAILIPASD ACKNVTLHPV SKLDAEKLVG RVNLKECFPTA 60

ANLIHSSDPD FQILEDGSVY TTNTILLSSE KRSPTILLSN TENQEKKKIF VFLEHQTKVL 120
KQRHTKEKVL RRAKRRWAPI PCSMLENSLG PFFLPFQQVQ SDTAQNYTII YSIRGPGVDQ 180
EPRNLFYVER DTGNLYCTRP VDRQYQESFE IIAFATTPDG YTPELPLPLI IKIEDENDNY 240
PIFTEETYTF TIFENCRVGT TVGQVCATDK DEPDTHMTRL KYSIIIGQVPP SPTLFSMHPT 300
5 TGVIITTSQQ LDRELIDIKYQ LKIKVQDMDG QYFGLQTTST CIINIDVDND HLPFTFRTSY 360
VTSVEENTVD VEILRVTVED KDLVNTANWR ANYTILKQNE NGNFKIVTDA KTNBGLVCVV 420
KPLNVEERQQ MILQIGVVNE APFSREASPR SAMSTATVTV NVEDQDEGPE CNPPIQTVM 480
KENAEVGTTS NGYKAYDPET RSSSGIRYK LTDPGTWVTI DENTGSIKVF RSLDREAETI 540
10 KNGIYNITVL ASDQGRCTCT GTLGIILQDV NDNSPFIPKK TVIICKPTMS SAEIVAVDPD 600
EPIHGPPPDF SLESSTSEVQ RMWRKKAIND TAARLSYQND PPGSYVVPV TVRDRMGMS 660
VTSLDVTLCD CITENDCTHR VDPRIIGGGV QLKGWAILAI LLGIALLFPI LFTLVGASG 720
TSKQPKVIPD DLAQONLIVS NTEAPGDDKV YSANGPTTQT VGASAQGVCG TVGSGIKNGG 780
QETIEMVKGQ HQTSSECRGA GHHTLDSR GGHTEVDNCR YTYSEWHSFT QPRLGEESIR 840
GHTLIKN

Seq ID NO: 620 DNA sequence
Nucleic Acid Accession #: NM_032545.1
Coding sequence: 46..718

1 11 21 31 41 51
AAACTGATCT TCAATGCACT AAGAGAAGGA GACTCTCAAA CCAAAAATGA CCTGGAGGCA 60
CCATGTCAAG CTTCGTGTTA CGGTCACTTT GGCATTACAG ATCATCAATT TGGGAAACAG 120
25 CTATCAAAGA GAGAAACATA ACGGCGGTAG AGAGGAAGTC ACCAAGTTG CCACTCAGAA 180
GCACGACGAG TCACCGCTCA ACTGGACCTC CAGTCATTTT GGAGAGGTGA CTGGGAGCGC 240
CGAGGGCTGG GGGCGGAGG AGCCGCTCCC TACTCCCGG GCTTTGCGAG AGGCTGCGTC 300
CGCGCGGGCG CGCTGCTGCA GGAACGGCGG TACCTGCGTG CTGGGCAAGT TCTGCGTGTG 360
CCGCGCCAC TTACCGCGC GCTACTGCGA GCATGACCA AGGCGCAGTG AATCGCGCGC 420
CCTGGAGCAC GGAGCCTGGA CCTTCCGCGC CTGCCACCTC TGCAAGTGCA TCTTCGGGCG 480
30 CCTGCACTGC CTCCCGCTTC AGACGCTTGA CCGCTGTGAC CCGAAAGACT TCCCTGCGCTC 540
CCACGCTCAC GGGCGGAGCG CGGGGGGCGC GCCACGCTG CTAATCTTGC TGCCCTGCGC 600
ACTCCTGCTC CGCTCTCTGC GCGCGGATGC GCGCGCGCAC CTGCGGTCCC TGGTCCCTTC 660
CGTCTCCAG CGGAGCGCGC GCGCTGCGG AAGCGCGGGA CTGGGCGATC GCCTTTAATT 720
35 TTCTATGTTG TAAATAATAG ATGTGTTTAG TTTACCGTAA GCTGAAGCAC TGGGTGAATA 780
TTTTTATTGG GTAATAAATA TTTTCATGAA AGCGCCAAAA AAAAAAAAAA AAAAAAAAAA 840
AAAAAA

Seq ID NO: 621 Protein sequence
Protein Accession #: NP_115934.1

1 11 21 31 41 51
MTWRHHVRL FTVSLALQII NLGNSVOREK HNGGREEVTK VATQKHRQSP LNWTSSEHFE 60
45 VTGSABGMGP EEPLPYSRAP GEGASARPRC CRNGGTGCVLG SPCVCPAHFT GRYCEHDQRR 120
SECGALEHGA WTLRAACHLCR CIFGALHCLP LQTPDRCDPK DFLASHAHGP SAGGAPSLLL 180
LLPCALLHRL LRPDAPAFHR SLVPSVLQRE RRPCGRPGLG HRL

Seq ID NO: 622 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..390

1 11 21 31 41 51
ATGAGGTTCAG GTGTCTCAGG CATGAGGACC GACTACCCCA GGAGTGTGCT GGCTCCTGCT 60
55 TATGTGTGAG TCTGTCTCCT CCTCTGTGT CCAAGGGAAG TCATCGCTCC CGCTGGCTCA 120
GAACCATGCG TGTGCCAGCC GGCACCCAGG TGTGGAGACA AGATCTACAA CCCCTTGGAG 180
CAGTGTCTGT ACAATGACG CATCGTGTCC CTGAGCGAGA CCGGCCAATG TGGTCCCCCC 240
TGCACTTCTT GCGCTGCTT TGAGCTCTGC TGTCTTGATT CCTTTGGCCT CACAAACGAT 300
60 TTTGTTGAGA AGCTGAAGGT TCAGGGTGTG AATTCACAGT GCCACTCATC TCCCATCTCC 360
AGTAAATGTG AAGAGGCGCG GATATGTTAG

Seq ID NO: 623 Protein sequence
Protein Accession #: FGENESH predicted

1 11 21 31 41 51
MRFSVSGMRT DYPRSVLAPA YVSVCLLLLC PREVIAPAGS EPWLQCPAPR CGDKIYNPLE 60
70 QCCYNDIAVS LSETRQCGPP CTFWPCFELC CLDSFGLTND FVVKLVQGV NSQCHSSPIS 120
SKCERGRIC

Seq ID NO: 624 DNA sequence
Nucleic Acid Accession #: M18728.1
Coding sequence: 51..1085

1 11 21 31 41 51
GGAGCTCAAG CTCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60
CCTCAGCCCC TCCTGCGAGA TTGCATGTCC CCTGGAAGGA GGTCTGCTC ACAGCCTCAC 120
80 TTCTAACCTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180
ATGTGCGAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAA CCGCCCCAG AATCGTATTG 240
GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATGTGA GGATATGTAA 300
TAGGAATCTA ACAAGCTACC CCAGGCGCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360
ATGCATCCCT GCTGATCCAG AACGTCAACC AGAATGACAC AGGATTCTAT ACCCTACAAG 420
85 TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGAGC 480
TGCCCAAGCC CTCCATCTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540
CCTTCACCTG TGAACCTGAG GTTCAGAA CAACCTACCT GTGGTGGGTA AATGGTCAGA 600
GCCTCCCGGT CAGTCCCGAG CTGCAGCTGT CCAATGGCAA CATGACCCTC ACTCTACTCA 660

5
 10
 15
 20
 25
 30

```

CGCTCAAAAG GAACGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720
ACCGCAGTGA CCCAGTACCC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCTC 780
CCTCAAAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCTGCCACC GCAGCCTCTA 840
ACCCACCTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900
TCTTTATCCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960
CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGCTCCTG 1020
TCCTCTCAGC TGTGGCCACC GTGGGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080
TATAGCAGCC CTGGGTATTT TTGGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCCT 1140
GAATCTCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGCTCTG 1200
CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAACCT AATGAAAATT TAAAGGGAAA 1260
ACCCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACCTAGAGA CAGTCAAACT 1320
GCAAAACCATG GTGAGAAATT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380
AACAAAGACTC CTCATCATGA TAAGGCTCTT ACCCCCTTTT AATTGTCTCT TGCTTATGCC 1440
TGCCCTCTTC GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500
GGGTAACTTA ACAGAGTGTC AGATCTATCT TGTCAATCCC AACGTTTTAC ATAAAAAAG 1560
AGATCCTTTA GTGCACCCAG TGACTGACAT TAGCAGCATC TTTAACACAG CCGTGTGTTT 1620
AAATGTACAG TGTGCTCTTT CAGAGTTGGA CTTCTAGACT CACCTGTTCT CACTCCCTGT 1680
TTTAATTCAA CCCAGCATG CAATGCCAAA TAATAGAATT GCTCCCTACC AGCTGAACAG 1740
GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800
GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860
CTGACTCATT CTTTATTCTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920
CTCTTGGTAT TACCCTCTTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980
CTCTAAAAGC TTTAAATGTC TGATGCAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040
GGCTGGAATT ACAAACCTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGS 2100
ATAAAGGCC CAATGTGGTG TAACGTATAA TAGCACTAAT GCTTTAAGAT TTGCTCACAC 2160
TCTCACCTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CACTGAAAT 2220
GTTAAGGAG AGAGTAGATC CAATTAAAAA AAATTAAAAA CAATTAAAAA AAAAAAAG 2280
ACACAGGAGA TTCCAGTCTA CTTGAGTTAG CATAATACAG AAGTCCCTCT TACTTTAACT 2340
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCAA TGTATTATT TCTGTGGTTC 2400
TGTTCCTTGG TTCCAAATTG ACAAACCCA CTGTTCTGTG ATTGTATTGC CCAGGGGAG 2460
CTATCACTGT ACTTGTAGAG TGGTGTGCT TTAATTCATA AATCAAAAT AAAAGCCAAT 2520
TAGCTCTATA ACT
  
```

Seq ID NO: 625 Protein sequence
 Protein Accession #: AA59907.1

35
 40
 45

```

1 11 21 31 41 51
MGPPSAPPCR LHPVKEVLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE VLLLAHNLPO 60
NRIGYSWYKG ERVDGNSLIV GYVIGTQQAT PGPAYSGRET IYPNASLLIQ NVTQNDTGPF 120
TLQVIKSLDLV NEEATQGFHV YPELPKPSIS SNNSNPVEDK DAVAPTCEPE VQNTTYLWNV 180
NGQSLVPSFR LQLSNGMMLT LLSVKRNDL GSYECIIONP ASANRSDPVT LNVLYGPDVP 240
TISPSKANYR PGENLNLSC AASNPPAQYS WFINGTFQQS TQELFIPNIT VNNSGSYMOC 300
AHNSATGLNR TTVTMITVSG SAPVLSAVAT VGITIGVLAR VALI
  
```

Seq ID NO: 626 DNA sequence
 Nucleic Acid Accession #: M18728.1
 Coding sequence: 1355..1657

50
 55
 60
 65
 70
 75
 80
 85

```

1 11 21 31 41 51
GGAGCTCAAG CTCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60
CCTCAGCCCC TCCTGTCAGA TTGCATGTCC CCTGGAAGGA GGTCTGCTC ACAGCCTCAC 120
TTCTAACCTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180
ATGTGCGAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAAA CCGCCCCAGG AATCGTATTG 240
GTTACAGCTG GTACAAAGCG GAAAGAGTGG ATGGCAACAG TCTAATGTGA GGATATGTAA 300
TAGGAACCTA ACAAGCTACC CCAGGGCCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360
ATGCATCCCT GCTGATCCAG AAGCTCACCC AGAATGACAC AGGATTCTAT ACCCTACAAG 420
TCATAAAGTC AGATCTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGAGAC 480
TGCCCAAGCC CTCCATCTCC AGCAACAACCT CCAACCCCGT GGAGGACAAG GATGCTGTGG 540
CCTTCACTGT TGAACCTGAG GTTCAGAAAC CAACCTACCT GTGGTGGGTA AATGGTCAGA 600
GCCCTCCCGT CAGTCCAGG CTGCAGCTGT CCAATGGCAA CATGACCCTC ACTCTACTCA 660
CGCTCAAAAG GAACGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720
ACCGCAGTGA CCCAGTACCC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCTC 780
CCTCAAAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCTGCCACC GCAGCCTCTA 840
ACCCACCTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900
TCTTTATCCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960
CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGCTCCTG 1020
TCCTCTCAGC TGTGGCCACC GTGGGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080
TATAGCAGCC CTGGGTATTT TTGGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCCT 1140
GAATCTCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGCTCTG 1200
CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAACCT AATGAAAATT TAAAGGGAAA 1260
ACCCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACCTAGAGA CAGTCAAACT 1320
GCAAAACCATG GTGAGAAATT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380
AACAAAGACTC CTCATCATGA TAAGGCTCTT ACCCCCTTTT AATTGTCTCT TGCTTATGCC 1440
TGCCCTCTTC GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500
GGGTAACTTA ACAGAGTGTC AGATCTATCT TGTCAATCCC AACGTTTTAC ATAAAAAAG 1560
AGATCCTTTA GTGCACCCAG TGACTGACAT TAGCAGCATC TTTAACACAG CCGTGTGTTT 1620
AAATGTACAG TGTGCTCTTT CAGAGTTGGA CTTCTAGACT CACCTGTTCT CACTCCCTGT 1680
TTTAATTCAA CCCAGCATG CAATGCCAAA TAATAGAATT GCTCCCTACC AGCTGAACAG 1740
GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800
GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860
CTGACTCATT CTTTATTCTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920
CTCTTGGTAT TACCCTCTTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980
CTCTAAAAGC TTTAAATGTC TGATGCAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040
GGCTGGAATT ACAAACCTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGS 2100
ATAAAGGCC CAATGTGGTG TAACGTATAA TAGCACTAAT GCTTTAAGAT TTGCTCACAC 2160
  
```

5
10
15
20
25
30
35
40
45
50
55
60
65

TCTCACCTAG GTGAGCSCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220
GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAAC CAATTTAAAA AAAAAAAGA 2280
ACACAGGAGA TTCCAGTCTA CTGAGTTAG CATAATACAG AAGTCCCTC TACTTTAACT 2340
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTTAACCA TGTATTATT TCTGTGGTC 2400
TGTTCTCTG TTCCAATTG ACAAAACCA CTGTTCTGT ATTGTATTG CCAGGGGAG 2460
CTATCACTGT ACTTGTAGAG TGGTGTGCT TTAATTCATA AATCACAAT AAAAGCCAAT 2520
TAGCTCTATA ACT

Seq ID NO: 627 Protein sequence
Protein Accession #: AAA59908.1

1 11 21 31 41 51
MDSFSQDVKT RLLIMIRLLP PFNLSLLMPA SPAWQDDAVI SISQEVASEG NLTECQIYLV 60
NPNVLHKIRD PLVHPVTDIS SIFNTAVCSN VQWSFSELDP

Seq ID NO: 628 DNA sequence
Nucleic Acid Accession #: M18728.1
Coding sequence: 2370..2501

1 11 21 31 41 51
GGAGCTCAAG CTCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60
CCTCAGCCCC TCCTCTCAGA TTGCATGTCC CCTGGAAGGA GGTCTGTCTC ACAGCCTCAC 120
TTCTAACCTT CTGGAACCCA CCCCACTCTC CCAAGCTCAC TATTGAATCC ACGCCATTCA 180
ATGTCGCGA GGGGAAGGAG GTTCTTCTAC TCGCCACAA CCTGCCCCAG AATCGTATTG 240
GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300
TAGGAATCTA ACAGCTACC CCAGGGCCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360
ATGCATCCCT GCTGATCCAG AACGTCAACC AGAATGACAC AGGATTCTAT ACCCTACAAG 420
TCATAAAGT AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGGAGC 480
TGCCCAAGCC CTCCATCTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540
CCTTCACTG TGAACCTGAG GTTCAGAAAC CAACCTACCT GTGGTGGGTA AATGGTCAGA 600
GCCTCCCGGT CAGTCCGAGG CTGCAGCTGT CCAATGGCAA CATGACCTCT ACTCTACTCA 660
GCGTCAAAAG GAACGATGCA GGATCTCTAT AATGTGAAAT ACAGAAACCA GCGAGTGCCA 720
ACCCAGTAGA CCCAGTCAAC CTGAATGTCC TCTATGGGCC AGATGTCCCC ACCATTTCCC 780
CCTCAAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA 840
ACCCACCTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900
TCITTTATCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960
CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGCTCCTG 1020
TCCTCTCAGC TGTGGCCACC GTGGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080
TATAGCAGCC CTGGTGATT TTGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCTT 1140
GAATTCCTCT AGCTCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGTCGT 1200
CTCTGCTCCT GAAGCCCTAT ATGCTGAGA TGGACAACCT AATGAAAATT TAAAGGGAAA 1260
ACCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACTAGAGA CAGTCAAACT 1320
GCAAAACATG GTGAGAAATT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380
AACCAAGACT CTGATCATGA TAAGGCTCTT ACCCCCTTTT AATTTGTCTT TGCTTATGCC 1440
TGCTCTTTC GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500
GGGTAACCTA ACAGAGTGTG AGATCTATCT TGTCATATCC AACGTTTAC ATAAAATAAG 1560
AGATCTCTTA GTGCACCCAG TGACTGACAT TAGCAGCATC TTAAACACAG CGTGTGTTC 1620
AAATGTACAG TGTGCTTTT CAGAGTTGGA CTTCTAGACT CACCTGTCTT CACTCCCTGT 1680
TTTAATTCAA CCCAGCCATG CAATGCCAAA TAATAGAAAT GCTCCCTACC AGCTGAACAG 1740
GGAGGAGTCT GTGCACTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800
GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860
CTGACTCATT CTTTATTCTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CBTAGTCCAA 1920
CTCTGGTAT TACCTCCTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980
CTCTAAAGC TTTAAATGTC TGCAATGAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040
GGCTGGAATT ACAAACTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100
ATAAAAGCCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTGGTCACAC 2160
TCTCACTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220
GTTAAGGAAG AAGATAGATC CAATTAAAAA AAATTAAAAA AAAAAAAGA 2280
ACACAGGAGA TTCCAGTCTA CTGAGTTAG CATAATACAG AAGTCCCTC TACTTTAACT 2340
TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTTAACCA TGTATTATT TCTGTGGTC 2400
TGTTCTCTG TTCCAATTG ACAAAACCA CTGTTCTGT ATTGTATTG CCAGGGGAG 2460
CTATCACTGT ACTTGTAGAG TGGTGTGCT TTAATTCATA AATCACAAT AAAAGCCAAT 2520
TAGCTCTATA ACT

Seq ID NO: 629 Protein sequence
Protein Accession #: AAA59909.1

1 11 21 31 41 51
MLTNVVISV LFPSCNLTKP TVLVLYCPGG AITVLVWCC PNS

Seq ID NO: 630 DNA sequence
Nucleic Acid Accession #: NM_016639.1
Coding sequence: 40..429

1 11 21 31 41 51
GCGGCGGCG CAGACAGCG CCGGCGCAGG ACGTGCACTA TGGCTCGGGG CTCGCTGCGC 60
CGGTTGCTGC GGTCTCTGTG GCTGGGCTC TGGCTGGCGT TGCTGCGCTC CBTGGCCGGG 120
GAGCAAGCGC CAGGACCCGC CCCCTGCTCC CGCGGAGCTC CCGGAGCGC GBACTGGAC 180
AAGTGATGG ACTGCGGCTC TTGAGGGCG CGACCGCACA GCGACTTCTG CCGTGGCTGC 240
GCTGAGCAC CTCCTGCCCC CTTCCGGCTG CTTTGGCCCA TCCTTGGGGG CGCTCTGAGC 300
CTGACCTTGG TGCTGGGCT GCTTCTGGC TTTTGGCTT GGAGACGATG CCGCAGGAGA 360
GAGAAGTCA CCACCCCAT AGAGGAGACC GCGCGAGAGG GCTGCCAGC TGTGGCGCTG 420

5

10

ATCCAGTGAC AATGTGCCCC CTGCCAGCCG GGGCTCGCCC ACTCATCATT CATTCAATCCA 480
 TTCTAGAGCC AGTCTCTGCG TCCAGAGCCG GCGGGGAGCC AAGCTCCTCC AACCAACAAGG 540
 GGGGTGGGGG GCGGTGAATC ACCTCTGAGG CCTGGGCCCA GGCTTCAGGG GAACCTTCCA 600
 AGGTGTCTGG TTGCCCTGCC TCTGGCTCCA GAACAGAAAG GGAGCCTCAC GCTGGCTCAC 660
 ACAAAACAGC TGACACTGAC TAAGGAACTG CAGCATTTGC ACAGGGGAGG GGGGTGCCCT 720
 CCTTCTTAG GACCTGGGGG CCAGGCTGAC TTGGGGGGCA GACTTGACAC TAGGCCCCAC 780
 TCACTCAGAT GTCTCGAAT TCCACCAGG GGGTCACCTT GGGGGGTAG GGACCTATT 840
 TTAACACTAG GGGCTGGCCC ACTAGGAGGG CTGGCCCTAA GATACAGACC CCCCCTACTC 900
 CCAAAAGCGG GGAGGAGATA TTTATTTTGG GAGGAGTTTG GAGGGGAGG AGAATTTATT 960
 AATAAAAGAA TCCTTAACCT TAAAAAATAA AAAAAAATAA

Seq ID NO: 631 Protein sequence
 Protein Accession #: NP_057723.1

15

20

1 11 21 31 41 51
 MARGSLRRL RLLVLGLWLA LLRSVAGEQA PGTAPCSRGS SWSADLDKCM DCASCRARPE 60
 SDFCLGCAA PPAPFRLLWP ILGGALSLTF VLGLLSGFLV WRRCRREKEF TPIETETGGE 120
 GCPAVALIQ

Seq ID NO: 632 DNA sequence
 Nucleic Acid Accession #: NM_003816.1
 Coding sequence: 79..2538

25

30

35

40

45

50

55

60

65

70

75

80

85

1 11 21 31 41 51
 CGGCAGGGTT GGAAATATGAT GGAAGAGGCG GAGGTGGAGG CGACCGAGTG CTGAGAGGAA 60
 CCTGCGGAAT CGGCCGAGAT GGGGTCTGGC GCGCGCTTTC CCTCGGGGAC CCTTCGTGTC 120
 CGGTGGTGGC TGTGTCTTGG CCTGGTGGGC CCAGTCTCTG GTGCGGCGCG GCCAGGCTTT 180
 CAACAGACCT CACATCTTTC TTCTTATGAA ATTATAACTC CTGGGAGATT AACTAGAGAA 240
 AGAAGAGAAG CCCCTAGGCC CTATTCAAAA CAAGTATCTT ATGTTATTCA GGCTGAAGGA 300
 AAGAGCATA TTATTCACTT GGAAAGGAAC AAAGACCTTT TGCCCTGAAGA TTTTGTGGTT 360
 TATACTTACA ACAAGGAAGG GACTTTAATC ACTGACCATC CCAATATACA GAATCAITGT 420
 CATTATCGGG GCTATGTGGA GGGAGTTTAT AATTCAATCCA TTGCTCTTAG CGACTGTTTT 480
 GGACTCAGAG GATGTCTGCA TTTAGAGAAAT GCGAGTTATG GGATTGAACC CCTGCAGAAC 540
 AGCTCTCATT TTGAGCACAT CATTATCGA ATGGATGATG TCTACAAAGA GCCTCTGAAA 600
 TGTGGAGTTT CCAACAAGGA TATAGAGAAA GAAACTGCAA AGGATGAAGA GGAAGAGCCT 660
 CCCAGCATGA CTCAGCTACT TCGAAGAAGA AGAGCTGTCT TGCCACAGAC CCGTATGTGT 720
 GAGCTGTCCA TTGTCTGAGA CAAGGAAAGG TATGACATGA TGGGAAGAAA TCAGACTGCT 780
 GTGAGAGAAG AGATGATTTCT CCTGGCAAAC TACTTGGATA GTATGTATAT TATGTTAAAT 840
 ATTCTGAATT TGCTAGTTGG ACTGGAGATT TGGACCAATG GAAACCTGAT CAACATAGTT 900
 GGGGTGTCTG GTGATGTGCT GGGGAACCTC GTGCAGTGGC GGGAAAGATT TCTTATCACA 960
 CGTCCGAGAC ATGACAGTGC ACAGCTAGTT CTAAGAAAGG GTTTTGTGTG AACTGCAGGA 1020
 ATGGCATTTG TGGGAACAGT GTGTTCAAGG AGCCACGCGG GCGGGATTAA TGTGTTTGGG 1080
 CAATCACTG GATGTCTGCA TGCTTCCATT GTTGCTCATG AATTGGGTCA TAATCTTGGG 1140
 ATGAATCAGC ATGATGGGAG AGATTGTTCC TGTGGAGCAA AGAGCTGCAT CATGAATCCA 1200
 GGAGCATCGG GTTCCAGAAA CTTTAGCAGT TGCACTGCAG AGGACTTTGA GAAGTTAACT 1260
 TTAATAAAG GAGGAAACTG CCTTCTTAAT ATTCCAAAGC CTGATGAAGC CTATAGTGCT 1320
 CCCCTCTGTG GTAATTAAGT GGTGGACGCT GGGGAAGAGT GTGACTGTGG TACTCCAAAG 1380
 GAATGTGAAT TGGACCTCTT CTGCGAAGGA AGTACCTGTA AGCTTAAATC ATTTGCTGAG 1440
 TGTGCTATG GTGACTGTGG TAAAGACTGT CGGTTCCTTC CAGGAGGTAC TTTATGCCGA 1500
 GGAATAACCA GTGAGTGTGA TGTTCAGAG TACTGCAATG GTTCTTCTCA GTTCTGTGAG 1560
 CCAGATGTTT TTATTCAGAA TGGATATCCT TGCCAGAAAT ACAAGGCTTA TTGCTACAAC 1620
 GGATCTGGCC AGTATTATGA TGCTCAATGT CAAGTCACTT TTGCTCATA AGCCAGAGCT 1680
 GCCCCCAAG AGTGTCTTCT TGAAGTGAAT TCTAAAGGTG ACAGATTGAG CAATTGTGGT 1740
 TTCTCTGGCA ATGAATACAA GAAGTGTGCC ACTGGGAATG CTTTGTGTGG AAAGCTTCAG 1800
 TGTGAGAAAT TACAAGAGAT ACCTGTATTG GGAATTGTGC CTGCTATTAT TCAACGCGCT 1860
 AGTCAGAGCA CCAATGTGTG GGGGTGTGGT TTCCAGCTAG GATCAGATGT TCCAGATCCT 1920
 GGGATGGTTA ACGAAGGCCA AAAATGTGGT GCTGGAAGAA TCTGTAGAAA CTTCAGATGT 1980
 GTAGATGCTT CTGTCTTGAA TTATGACTGT GATGTTTCTA AAAAGTGTCA TGGACATGGG 2040
 GTATGTAATA GCAATAGAAA TTGTCACTGT GAAATGGCT GGGCTCCCCC AAATGTGTAG 2100
 ACTAAGGAT ACGGAGGAAG TGTGGACAGT GGACCTACAT ACAATGAAT GAATACGTCA 2160
 TTGAGGGACG GACTTCTGGT CTTCTCTTC CTAATTGTTT CCCTTATTGT CTGTGCTATT 2220
 TTTATCTTCA TCAAGAGGGA TCAACTGTGG AGAAGCTACT TCAGAAAGAA GAGATCACA 2280
 ACATAGAGT CAGATGGCAA AAATCAAGCA AACCCTTCTA GACAGCCGGG GAGTGTTCCT 2340
 CGACATGTTT CTCAGTGAC ACCTCCAGCA GAAGTTCCTA TATATGCAA CAGATTGCA 2400
 GTACCAACCT ATGCAGCAA GCAACCTCAG CAGTTCCTAT CAAGGCCACC TCCACCAAA 2460
 CCGAAAGTAT CATCTCAGG AAACTTAATT CCTGCCGTC CTGCTCTGTC ACCTCCTTTA 2520
 TATAGTTCCC TCACTTGATT TTTTAACTT TCTTTTGTCA AATGCTTCTA GGGAACTGAG 2580
 CTAATACTTT TTTTCTTCT TGATGTTTTC TTGAAAAGCC TTTCTGTGTC AACTATGAAT 2640
 GAAACAAAA CACCACAAA CAGACTTCAC TAACACAGAA AAACAGAAAC TGAGTGTGAG 2700
 AGTTGTGAAA TACAAGGAAA TGCAGTAAAG CCAGGGAATT TACAATAACA TTTCCGTTTC 2760
 CATCATGAAA TAAGTCTTAT TCAGTCACTG GTGAGGTTAA TGCATTAAT ATGGATTITT 2820
 TGAACATGTT ATTGCACTGA TTCTCAAAAT AACTGTATTG GTGTAAGATT TTTGTCTATT 2880
 AGTGTTAAG TGTATTCTG AATTCTCTAC CTTAGTTATC ATTAATGTAG TTCCTCATG 2940
 AACATGTGAT AATCTAATAC CTGTGAAAAC TGACTAATCA GCTGCCAATA ATATCTAATA 3000
 TTTTTCATCA TGCACGAATT AATAATCATC ATACTCTAGA ATCTGTCTG TCACTCACTA 3060
 CATGAATAAG AAGCAAGGAC CTTCAAAAGA ATGCACAGA ACCACAATA AGATGTCATA 3120
 TTTATTTGAA AGTACAAAT ATACTAAAAG AGTGTGTGTG TATTCACGCA GTTACTCGCT 3180
 TCCATTTTTA TGACCTTTCA ACTATAGGTA ATAACCTCTA GAGAAATTA TTTAATATTA 3240
 GAATTTCTAT TATGAATCAT GTGAAAGCAT GACATTGTT CACAATAGCA CTAATTTAAA 3300
 TAAATTATAA GCTTTAAGGT ACGAAGTATT TAATAGATCT AATCAAATAT GTTGATTCT 3360
 GGCTATAATA AAGCAGGAGC AATTATAAAA TCTTCAATCA ATTGAACCTT TACAAAACCA 3420
 CTTGAGAAAT TCAATGAGCA TTTAAATCT GAACCTTCAA AGCTTGCTAT TAAATCATTT 3480
 AGAATGTTTA TCACTACTAA GGTGTGCTGG GTCATGTAAA ATATTAGACA CTAATATTTT 3540
 CATAGAAATT AGGCTGGAGA AAGAAAGGAG AAATGGTTTT CTTAAATACC TACAAAAG 3600
 TTACTGTGGT ATCTATGAGT TATCATCTTA GCTGTGTTAA AAATGAATTT TTACTATGGC 3660

AGATATGGTA TGGATCGTAA AATTTTAAGC ACTAAAAATT TTTTCATAAC CTTTCATAAT 3720
AAAGTTTAAAT AATAGGTTTA TTAAGTGAAT TTCAATAGTT TTTTAAAAGT GTTTTGGTT 3780
TGTGTATATA TACATATACA AATACAACAT TTACAATAAA TAAATACTT GAAATTCTCA 3840
AAAAAAAAA AAAAAAAAAA AAAAA

Seq ID NO: 633 Protein sequence
Protein Accession #: NP_003807.1

1 11 21 31 41 51
MSSGARPPSG TLRVRWLLLL GLVGPVLGAA RPFQQTSHL SSYEIITPWR LTRERREAPR 60
PYSKQVSVVI QAEGKEHIIH LERNKDLLPE DFVVTYNKE GTLITDHPNI QNHCHYRGVY 120
EGVHNSIAL SDCFLGRLL HLENASYGIE PLQNSHFEH IYRMDVYK EPLKCGVSNK 180
DIEKETAKDE EEPFPMQTL LRRRAVLPO TRYVELFIV DKERYDMMGR NQAVREEMI 240
LLANYLDSMY IMLNIRIVLV GLEIWTNGNL INIVGGAGDV LGNFVQWREK FLITRRRHDS 300
AQLVLKKGFG GTAGMAFVGT VCSRSAGGI NVFQGITVET FASIVAEHLG HNLGMNHDDG 360
RDCSCGAKSC IMNSGASGSR NFSSCSAEDF EKLTLNKGKN CLLNIPKDE AYSAPSCGNK 420
LVDAGEEDCD GTPKECELP CCEGSTCKLK SFAECAYGDC CKDCRFLPGG TLRCKGTSEC 480
DVPEYCNSSG QFCQPDVFIQ NGYPCQNNKA YCYNGMCQY DAQCQVIFGS KAKAAPKDCP 540
IEVNSKGRPF GNCGPSGNEY KKCATGNALC GKLCQENVQE IPVFGIVPAI IQTPSRGTCK 600
WGVDFQLGSD VPPFGMVNKG TKOGAGKICR NFQCVDAVL NYDCDVQKCK HGHGVCSNKN 660
NHCENGMAP FNCETKGYGG SVDSGPTYNE MNTALRDGLL VFFFLIVPLI VCAIFIFIKR 720
DQLWRSYFRK KRSQYESDG KNOANPSRQP GSVPRHVSVP TPPREVPIYA NRFVPTTAA 780
KQPQKPPSRP PPPQPKVSSQ GNLIIPARPAP APPLYSLLT

Seq ID NO: 634 DNA sequence
Nucleic Acid Accession #: NM_002091.1
Coding sequence: 56..503

1 11 21 31 41 51
AGTCTCTGCT CTTCCAGGCC TCTCCGGCGC GCTCCAAGGG CTTCCCGTCG GGACCATGCG 60
CGGCGAGTGA CTCGCCGTGG TCTGCTGGCC GCTGGTCCTC TGCCCTAGCGC CCCGGGGGCG 120
AGCGGTCCCG CTGCTGCGGG GCGGAGGGAC CGTGCTGACC AAGATGTACC CGCGCGGCAA 180
CCACTGGGCG GTGGGCACT TAATGGGGAA AAAGAGCACA GGGGAGTCTT CTTCTGTTTC 240
TGAGAGAGGG AGCTGAAGC AGCAGCTGAG AGAGTACATC AGGTGGGAAG AAGCTGCAAG 300
GAATTTCTGCT GGTCTCATAG AAGCAAAGGA GAACAGAAAC CACCAGCCAC CTCACCCCAA 360
GGCCTTGGGC AATCAGCAGC CTCTGTGGGA TTCAGAGGAT AGCAGCAACT TCAAGATGT 420
AGGTTCAAAA GGCAGAGTTG GTAGACTCTC TGCTCCAGGT TCTCAACGTG AAGGAAGGAA 480
CCCCAGCTG AACCAAGCAAT GATAATGATG GCCTCTCTCA AAAGAGAAAA ACAAACCCCC 540
TAAGAGACTG AGTCTGCAAA GCATCAGTTC TACGGATCAT CAACAAGATT TCCTTGTGCA 600
AAATATTCTA CTATTCTGTA TCTTTTATCC TTGACTAAAT TCGTGATTTC CAAGCAGCAT 660
CTTCTGGTTT AAACCTTGTG GCTGTGAACA ATTGTGCAAA AGAGTCTTCC AATTATAGCT 720
TTTTTATATC TAGGCTACCT GTTGGTTAGA TTCAAGGCC CGAGCTGTTA CCATTACAAA 780
TAAAGCTTA AACACAT

Seq ID NO: 635 Protein sequence
Protein Accession #: NP_002082.1

1 11 21 31 41 51
MRGSELPLVL LALVLCAPR GRAVPLPAGG GTVLTKMYPR GNEHVAVGHM GKSTGESSS 60
VSRGSLKQQ LREYIRWEA ARNLLGLIEA KENRNHQPQP PKALGNQQPS WDSSESSNFK 120
DVSGSKGVGR LSAPGQREG RNPQLNQQ

Seq ID NO: 636 DNA sequence
Nucleic Acid Accession #: NM_016522.1
Coding sequence: 265..1299

1 11 21 31 41 51
GCGGAAGCAG CGAGGAGGGA GCGCCCTTTG GCGGTCCTCC GTGGAACCGG TTTTCCGAGG 60
CTGGCAAAAG CGAGGCTGG ATTTGGGGGA GGAATATTAG ACTCGAGGGA GTCTGCGCGC 120
TTTTCTCCTC CCGCGCCTC CCGGTGCGCG CGGGTTCAAC GCTCAGTCCC CGCGCTCGCT 180
CCGCAACCCA CCCACTTCCT GTGCTCGCCC GGGGGGCGTG TGCCGTGCGG CTGCGGAGT 240
TOGGGGAAGT TGTGGCTGTC GAGAATGGGG GTCTGTGGGT ACCTGTTCCT GCCCTGGAAG 300
TGCCCTGCTG TCGTGTCTCT CAGGCTGCTG TTCCTTTGAT CCACAGGAGT GCCCGTGCGC 360
AGCGGAGATG CCACCTTCCC CAAGCTATG GACAAAGTGA CGGTCCGGCA GGGGAGAGC 420
GCCACCTCA GGTGCACTAT TGACAAACCG GTACCCCGGG TGCCCTGGCT AAACCGCAGC 480
ACCATCTCT ATGCTGGGAA TGACAAGTGG TGCCCTGGATC CTGCGTGGT CTTCTGAGC 540
AACACCCAAA CGCAGTACAG CATCGAGATC CAGAACGTGG ATGTGTATGA CGAGGGCCCT 600
TACACCTGCT CGGTGAGAGC AGACAACAC CCAAGACCT CTAGGGTCCA CCTCATTTGT 660
CAAGTATCTC CCAAAATTTG AGAGATTCT TCAGATATCT CCATTAATGA AGGGAACAAT 720
ATTAGCTCA CCGTCAATAG AACTGGTAGA CCAGAGCCTA CGGTACTTGA GAGACATC 780
TCTCCCAAAG CGGTGGCTT TGTGAGTGA GACGAATACT TGGAAATTCA GGGCATCACC 840
CGGGAACAGT CAGGGGACTA CGAGTGCAGT GCCTCCAATG AGTGGCCCG GCCCGTGGTA 900
CGGAGAGTAA AGGTACCGGT GAACATATCA CCATACATTT CAGAGCCAA GGGTACAGT 960
GTCCCGTGG GACAAAGGG GACACTGCAG TGTGAAGCCT CAGCAGTCCC CTCAGCAGAA 1020
TTCCAGTGGT ACAAGATGA CAAAGACTG ATTGAAGGAA AGAAGGGGT GAAAGTGGAA 1080
AACAGACTT TCTCTCAAA ACTCATCTTC TTCATGTCT CTGAACATGA CTATGGGAAC 1140
TACACTTGG TGGCTCCAA CAAGCTGGC CACACCAATG CCAGCATCAT GCTATTGTGT 1200
CCAGGCGCG CTAGCGAGGT GAGCAACGCG ACGTCGAGGA GGGCAGGCTG GCTCTGGCTG 1260
CTGCTCTTC TGGTCTTGA CCGCTTCTC AAATTTTGT GTGAGTGCCA CTTCCCAACC 1320
CGGGAAGGAT TGGCGCACCC ACCACCAACA ACACAACAGC AATGGCAACA CGACAGCAA 1380
CCAATCAGT ATATACAAAT GAAATTAGAA GAAACAAGC CTCATGGGAG AGAAATTGTA 1440
GGGAGGGGAA CAAAGATATC TTTGGGGGGA AAAGAGTTTT AAAAAAGAAA TTGAAATTTG 1500
CCTTCAGAT ATTTAGGTAC AATGGAGTTT TCTTTTCCCA AACGGGAAGA ACACAGCACA 1560

CCCGGCTTGG ACCCACTGCA AGCTGCATCG TGCAACCTCT TTGGTGCCAG TGTGGGCAAG 1620
 GGCTCAGCCT CTCTGCCAC AGACTGCCCC CACGTGGAAC ATTCCTGGAGC TGGCCATCCC 1680
 AAATTCAATC AGTCCATAGA GACGAACAGA ATGAGACCTT CCGGCCCAAG CGTGGCGCTT 1740
 CGGGCCCAAG CGTGGCGCTG CGGGCACTTT GGTAGACTGT GCCACCACGG CGTGTGTGTG 1800
 GAAACGTGAA ATAAAGAG CAAAAA AAAAAA

Seq ID NO: 637 Protein sequence
 Protein Accession #: NP_057606.1

1 11 21 31 41 51
 MGVCYGLFLP WKCLVVVSLR LLFLVPTGVP VRSGDATFPK AMDNVTVRQG ESATLRCTID 60
 NRVTIRVANLN RSTILYAGND KWCLDFRVVL LSNQTQYSI EIQNVDVYDE GPYTCSVQTD 120
 NHPKTSRVHL IVQVSPKIVE ISSDISINEG NNISLTCIAT GRPEPTVTWR HISPKAAGVF 180
 SEDEYLEIQG ITRQSGDYE CSASNDVAAP VVRRVKVTN YPPYISEAKG TGVPVQKQGT 240
 LQCEASAVPS AEPQYKDDK RLIEGKKGVK VENRPFSLKL IFPNVSEHDY GNYTCVASNK 300
 LGHTNASIML FPGAVSEVS NGTSRRAGCV WLLPLLVLHL LLKF

Seq ID NO: 638 DNA sequence
 Nucleic Acid Accession #: NM_012261.1
 Coding sequence: 203..1045

1 11 21 31 41 51
 GATTTGCTCT GCCAGCAGCT GTGGGTGCGG CGCTCGACAC CGAGTCCTAG CTAGGCGCTC 60
 ACAGAAATAG CGCTCCCTCC CTCCCCTTTC TCTGTCCCC GCCTCTCGCT CACCCCGGCC 120
 CACTCCAGCG GCGACTTTGA GGGATTCCCT CTCTGGCGGC CTCTGCAGCA GCACAGCCGG 180
 CCTCATTGCG GGCAGTCGGA GTATGGATCT CCAAGGAAGA GGGGTCCCCA GCATCGACAG 240
 ACTTCGAGTT CTCTGATGT TGTTCATAC AATGGCTCAA ATCATGGCAG AACAGAAGT 300
 GGAAATCTC TCAGGCGCTT CCACTAACCC TGAAAAAGAT ATATTGTGG TCGGGGAAAA 360
 TGGGACGACG TGTCTCATGG CAGAGTTTGC AGCCAAATTT ATTGTACCTT ATGATGTGTG 420
 GGCCAGCAAC TACGTAGATC TGATCAGAGA ACAGGCCGAT ATCGATTGA CCGGGGAGC 480
 TGAGGTGAAG GGCCTGTGTG GCCACAGCCA GTCGGAGCTG CAAGTGTTCCT GGTGGATCG 540
 CGCATATGCA CTCAAATGTC TCTTTGTAAA GGAAGGCCAC AACATGTCCA AGGGACCTGA 600
 GGCGACTTGG AGGCTGAGCA AAGTGCAGTT TGTCTACGAC TCCTCGGAGA AAAACCACTT 660
 CAAAGACGCA GTCAGTGTCT GGAAGCACAC AGCCAACTCG CACCACCTCT CTGCTTGGT 720
 CACCCCGCT GGAAGTCTCT ATGAGTGTCA AGCTCAACAA ACCATTTCAC TGGCCTCTAG 780
 TGATCCGCGC AAGACGCTCA CCAATGATCCT GTCTGCCGTC CACATCCAAC CTTTTCAGAT 840
 TATCTCAGAT TTTGCTTCA GTGAAGAGCA TAAATGCCCA GTGGATGAGC GGGAGCAACT 900
 GGAAGAAACC TTTGCGCTGA TTTTGGGGCT CATCTTGGGC CTGCTCATCA TGGTAACACT 960
 CGGATTTAC CAGCTCCACC ACAAAATGAC TGCCAAACAG GTGCAGATCC CTCGGGACAG 1020
 ATCCAGATAT AAGCAGTAGG GCTAGAGGCC GTTAGGCAGG CACCCCTAT TCCTGCTCCC 1080
 CCAACTGGAT CAGGTAGAAC AAAAAAGCA CTTTTCATC TTGTACAGCA GATACACCA 1140
 CATAGCTACA ATCAACAGG CTTGGGTATC TGAGGCTTGC TTGGCTTGTG TCCATGCTTA 1200
 AACCCACGGA AGGGGAGAGC TCTTTCGGAT TTGTAGGGTG AAATGGCAAT TATTCTCTCC 1260
 ATGCTGGGGA GGAGGGGAGG AGGCTCTCAG ACAGCTTTCG TGCTCATGGT GGCTTGGCTT 1320
 TGACTCTCCA AAGAGCAATA AATGCCACTT GGAGCTGTAT CTGGCCCAAA AGTTTAGGGA 1380
 TTGAAACAT GCTTCTTTGA GGAGGAAACC CCTTTAGGTT CAGAAGAATA TGGGTGCTT 1440
 TGCTCCCTTG GACACAGCTG GCTTATCCTA TACAGTTGTC AATGCACACA GAATACAAAC 1500
 TCATGCTCCC TTGACCAAGA CCCCAGAAAG TGATTATGTC TTCTGGCTGG CATTCTGCAT 1560
 GTTTAGTATG TGTCTGGGA ATGTTTCACT GCTACCCGCA TCCAGCGACT GCACACCAAG 1620
 AAAACGACTA ATGTAATAT CAGAGTTTGT TTGGACTTCT TCCTGTGCCA GGTCCAAGTC 1680
 GGGGACCTG AAGAATCAAT CTGTGTGAGT CTGTTTTTCA AAATGAAATA AAACACACTA 1740
 TTCTCTGGC

Seq ID NO: 639 Protein sequence
 Protein Accession #: NP_036393.1

1 11 21 31 41 51
 MDLQGRGVPS IDRLRVLLML PHTMAQIMAE QEVENLSGLS TNPEKDIFVV RENGTTCLMA 60
 EFAAKFIPVY DVWASNYVDL ITSQADIALT RGAEVKGRGC HSQSELQVFW VDRAYALKML 120
 FVKESHNSMK GPEATWRLSK VQPVYDSSEK THFKDAVSAG KHTANSHLS ALVTPAGKSY 180
 EQQAQTISL ASSDPQKTVT MILSAVHIQF FDIISDFVPS ESHKCPVDER BQLEETLPLI 240
 LGLILGLVIM VTLAIYVHH KMTANQVQIP RDRSQYKHM

Seq ID NO: 640 DNA sequence
 Nucleic Acid Accession #: NM_002993.1
 Coding sequence: 64..408

1 11 21 31 41 51
 GGCACGAGCC AGTCTCCGGC CCTCCACCCA GCTCAGGAAC CCGGAAACCC TCTCTGAAC 60
 ACTATGAGCC TCCCGTCCAG CCGCGCGGCC CGTGTCCCGG GTCCTTCGGG CTCCTTGTGC 120
 GCGCTGCTCG CGCTGCTGCT CTTGCTGAGC CCGCCGGGGC CCCTCGCCAG CGCTGGTCTT 180
 GTCTCTGCTG TGCTGACAGA GCTGCGTTGC ACTTGTTTAC GCGTTACGCT GAGAGTAAC 240
 CCCAAACGGA TTGGTAACT GCAGGTGTTT CCGCGAGGCC CGCAGTGCTC CAAGGTGGAA 300
 GTGGTAGCCT CCGTGAAGAA CGGGAAGCAA GTTTGTCTGG ACCCGGAAGC CCCTTTTCTA 360
 AAGAAAGTCA TCCAGAAAT TTTGGACAGT GGAACCAAGA AAAACTGAGT AACAAAAAG 420
 ACCATGCATC ATAAATTTGC CCACTCTTCA GCGGAGCAGT TTTCTGGAGA TCCCTGGAAC 480
 CAGTAAGAAT AAGAAGGAAG GGTGGTTTTT TTTCCATTTT CTACATGAT TCCCTACTTT 540
 GAAGAGTGTG GGGGAAGGCC TACGCTTCTC CCTGAAGTTT ACAGCTCAGC TAATGAAGTA 600
 CTAATATAGT ATTTCACTA TTTACTGTTA TTTTACCTGA TAAGTTATTG AACCTTTGG 660
 CAATTGACCA TATTGTGAGC AAGAAATCAC TGGTTATTAG TCTTTCAATG AATATTGAAT 720
 TGAAGATAAC TATTGTATT CTATCATACA TTCCTTAAAG TCTTACCGAA AAGGCTGTGG 780
 ATTTCTGATG GAAATAATGT TTTATTAGTG TGCTGTTGAG GGAGGTATCC TGTGTTCTT 840
 ACTCACTCTT CTCATAAAT AGGAAATATT TTAGTTCTGT TTTCTTGGG AATATGTTAC 900

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

TCTTTACCT AGGATGCTAT TTAAGTTGTA CTGTATTAGA AACTGGGGTG TGTCTATCCG 960
TTATCTGTGC AGAATATATT TCCTTATTCA GAATTTCTAA AAATTTAAGT TCTGTAAGGG 1020
CTAATATATT CTCTTCCTAT GGTITTAGAT GTTGTAGTGC TTCTTAGTAT GGCATAATGT 1080
CATGATTTAC TCATTAAAT TTGATTTTGT ATGCTATTTT TTCACTATAG GATGACTATA 1140
ATTCTGGTCA CTAATATATAC ACTTTAGATA GATGAAGAAG CCCAAAAACA GATAAATTCC 1200
TGATGCTGTA TTACATAGA AATGTATTCT CTGTGGTTTT TAAATAAAG CAAAATTAAAC 1260
AATGATCTGT GCTCTGCAAA GTTTTGAAAA TATATTGAA CAATTTGAAT ATAAATTCAT 1320
CATTTAGTCC TCAAAATATA TACAGCATGT CTAAGATTTT CAGATATCTA TTGTGGATCT 1380
TTTAAAGGTT TTGACCATTT TGTATGAGG AATTATACAT GTATCACATT CACTATATTA 1440
AAATGCACT TTTATTTTTT CCTGTCTGTC ATGTTGGTTT TTGGTACTTG TATTGTCAAT 1500
TGGAGAAACA ATAAAGATT TCTAAACCAA AAAAAAAAAA AAAAAA

Seq ID NO: 641 Protein sequence

Protein Accession #: NP_002984.1

1 11 21 31 41 51
MSLPSSRAAR VPGPSGLCA LLALLLLLP FGPLASAGPV SAVLTELRCT CLRVTLRVNP 60
KTIGKLQVFP AGPQCSKVEV VASLKNQKV CLDPEAPFLK KVIQKILDSG NKKK

Seq ID NO: 642 DNA sequence

Nucleic Acid Accession #: NM_013271.1

Coding sequence: 27..809

1 11 21 31 41 51
TCCGGAGCCA GGCTCGCTGG GGCAGCATGG CGGGGTCGCG GCTGCTCTGG GGGCGCGGG 60
CGGGGGGGGT CGGCCCTTTTG GTGCTGCTGC TGCTCGGCCT GTTTCGGCCG CCCCCCGCG 120
TCTGCGCGCG GCCGTAAAG GAACCCCGCG GCCTAAGCGC AGCGTCTCCG CCCCTGGCTG 180
AGACTGGCGC TCCTCGCGCG TTCCGGCGGT CAGTGCCTCG AGGTGAGGCG GCGGGGGCGG 240
TGCAGGAGCT GCGCGCGGCG CTGGCGCATC TGCTGGAGGC CGAACGTCTAG GAGCGGCGCG 300
GGCGCGAGGC GCAGAGGCTG GAGGATCAGC AGGCGCGCGT CTTGGCGCAG CTGCTGCGCG 360
TCTGGGGCGC CCCCCGCAAC TCTGATCCCG CTCTGGCGCT GAGCGACGAC CCGGACGCGC 420
CTGCAGCGCA GCTCTGCTGC GCTCTGCTCC GCGCCGCGCT TGACCTGCGC GCCCTAGCAG 480
CCGAGCTTGT CCGCGCGCGC GTCCCGCGCG CGGCGCTCCG ACCCGCGCCC CCGTCTCTAG 540
ACGACGCGCC CCGCGCGCGC GATGCTGAGG AGGCAGGCGA CGAGACACCC GACGTGGACC 600
CGAGCTGTT GAGGTACTTG CTGGGACGGA TTCTTGCGGG AAGCGCGGAC TCCGAGGGGG 660
TGGCAGCCCC GCGCGCGCTC CCGCGTCCG CCGACCAAGA TGTGGCTCTT GAGCTGCCCC 720
CTGAGGGCGT GCTGGGGGCG CTGCTGCGTG TGAACGCGCT AGAGACCCCG GCGCCCCAGG 780
TGCCTGCAAG CCGCTCTTGT CCACCTGAG CACTGCGCGG ATCCCGTGCA CCTTGGGACC 840
CAGAAGTGCC CCGCGCATCC CGCCACCAGG ACTTCTCCCG GCCAGCACGT CCAGAGCAAC 900
TTACCCCGGC CAGCCAGCCC TCTCACCGA GGATCCCTAC CCCTGGGCC ACAATAACAT 960
GATCTGAGC

Seq ID NO: 643 Protein sequence

Protein Accession #: NP_037403.1

1 11 21 31 41 51
MAGSPLLWGP RAGGVGLLV LLLGLFRPPP ALCARPVKEP RGLSAASPPL AETGAPRRFR 60
RSVPGEAAG AVQELARALA HLLAEERQER ARAEAQEAED QOARVLAQLL RVWGAPRNSD 120
PALGLDDDDP APAAQLARAL LRRLDPAAL AAQLVPAPVP AAALRFRPVP YDDGAPGPD 180
EEAGDETPDV DPELLRYLLG RILAGSADSE GVAAPRRLRR AADHDVGSSEL PPEGVLGALL 240
RVKRLTTPAP QVPARRLLFP

Seq ID NO: 644 DNA sequence

Nucleic Acid Accession #: NM_002214

Coding sequence: 681..2990

1 11 21 31 41 51
CCCAGAGCCG CCTCCCCCTG TTGCTGGCAT CCCGAGCTTC CTCCTTGCC AGCCAGGAGC 60
CTGCCGACTT GTCTTTGGCC GCTGCTCCGC AGACGGGGCT GCRAAGCTGC AACTAATGTT 120
GTTGGCTCCG CTGCCACCT GTGGAAGCAA CTGCGCTGAT TGATGCGCCA CAGACTTTTT 180
TCCCCTGAGC CTCGCCGCGC TACCCTCCCA CAGATCCAGC ATCACCCAGT GAATGTACAT 240
TAGGGTGGTT TCCCCCAGC CTTGCGGCTT GTTTGGGTT TGATTGTGTT TGGCTCTTGG 300
CTAAGCTGAT TTATGAGCA GAAGCCCCAC CGGCTGAGGA GAAACAAAAG CTCCTTTCTT 360
TGTCCCGGAG CAGGCTGCGG AGCCCTTGCA GAGCCCTCTC TCCAGTCCGC GCCGGGCCCT 420
TGGCGTGA AGGAGGTGCT TCTCGCGGAG ACCCGGGGAC CGCGCTGCC GAGCGCGGAG 480
GGCGGTAGGG GCCCTGAGAT GCGGAGCGGT GCCCGGGCCC GCTTACCTGC ACCGCTTGCT 540
CCGAGCCGCG GGGTCCGCTT GCTAGGCTTG CGGAAAAGT CCTAGCGACA CTCGCCCGCG 600
GGCCCCGAGG TCGCCCGGGA GGCAGAGCCC GCGTCCGAA GGCAGCCAGC CGCGCGCGCG 660
GGGGGGGGCT GTTTTGCAAT ATGTGCGGCT CGGCCCTGGC TTTTITTTACC GCTGCATTTG 720
TCTGCTGCA AAACGACCGG CAGAGTCCCG CCTGTTCTCT CTGGGAGGCC TGGGTGTTTT 780
CACTTGTCT TGACTGGGC CAAGGTGAAG ACRATAGATG TGCACTTTCA AATGCAGCAT 840
CCTGTGCCAG GTGCTTGCG CTGGTCCAG AATGTGATG GTGTGTTCAA GAGGATTTC 900
TTTCAGGTGG ATCAAGAAAT GAACGTTGTT ATATTGTTT CAATTTAATA AGCAAAGGCT 960
GCTCAGTTGA TTCAATAGAA TACCATCTG TGCACTGTTT AATACCCACT GAAATGAGAA 1020
TAATATACCA GGTGACACCA GGAGAAGTGT CTATCCAGCT GGTTCAGGA GCGGAGGCTA 1080
ATTTTATGCT GAAAGTTTAT CCTCTGAAGA AATATCCTGT GGATCTTTAT TATCTGTTG 1140
ATGTCTCAGC ATCAATGCAC AATAATATAG AAAAATTAAA TTCGTTTGA AACGATTAT 1200
CTAGAAAAAT GGCATTTTTC TCCCGTGACT TTCGTTCTGG ATTTGGCTCA TACGTTGATA 1260
AAACAGTTTC ACCATACATT AGCATCCACC CCGAAAGGAT TCATAATCAA TGCACTGACT 1320
ACAAATTAGA CTGATGCCCT CCCCATGGAT ACATCCATGT GCTGTCTTTG ACAGAGAACA 1380
TCACTGAGTT TGAGAAGACA GTTCATAGAC AGAAGATCTC TGGAAACATA GATACACCAG 1440
AAGGAGGTTT TGACGCCATG CTTGAGGCG CTGTCTGTGA AAGTCATATC GGATGGCGAA 1500
AAGAGGCTAA AAGATTGCTG CTGGTGATGA CAGATCAGAC GTCTCATCTC GCTCTTGATA 1560

5
10
15
20
25
30
35
40

GCAAAATGGC AGGCATAGTG GTGCCCAATG ACGGAAACTG TCATCTGAAA AACCAACGCTCT 1620
ACGTCAAATC GACCAACCATG GAACACCCCT CACTAGGCCA ACTTTCAGAG AAATTAATAG 1680
ACAAACAATC TAATGTCAATC TTTGCAATTC AAGGAAAAACA ATTTCAATGG TATAAGGATC 1740
TTCTACCCCT CTGCGCAGGC ACCATTGCTG GTGAAATAGA ATCAAAGGCT GCAAAACCTCA 1800
ATAAATTTGGT AGTGGAGGCC TATCAGAAGC TCATTTCAGA AGTGAAAGTT CAGGTGGGAA 1860
ACCAGGTACA AGGCATCTAT TTTAAACATTA CCGCCATCTG TCCAGATGGG TCCAGAAAGC 1920
CAGGCGATGA AGGATGACGA AACGTGAAGA GCAATGATGA AGTTCTTTTC AATGTAAACG 1980
TTACAATGAA AAAATGTGAT GTCAACAGGAG GAAAAAATA TGCAATAATC AAACCTATTG 2040
GTTTAAATGA AACCGCTAAA ATTCATATAC ACAGAACTG CAGCTGTGAG TGTGAGGACA 2100
ACAGAGGACC TAAAGGAAAG TGTGTAGATG AAACCTTTCT AGATTCCAAG TGTTTCCAGT 2160
GTGATGAGAA TAAATGTCTAT TTTGATGAAG ATCAGTTTTC TTCTGAGAGT TGCAAGTCAC 2220
ACAAAGATCA GCGTGTGTC AGTGGTGGAG GAGTTTGTGT TTGTGGGAAA TGTTCATGTC 2280
ACAAATATAA GCTTGGAAAA GTGTATGGAA AATACTGTGA AAAGGATGAC TTTTCTGTCT 2340
CATATCACA TGGAAATCTG TGTGCTGGGC ATGGAGAGTG TGAAGCAGGC AGATGCCAAT 2400
GCTTCAGTGG CTGGGAAGGT GATCGATGCC AGTGCCCTTC AGCAGCAGCC CAGCACTGTG 2460
TCAATTCAAA GGGCCAAGTG TGCAAGTGGAA GAGGCACGTG TGTGTGTGGA AGGTGTGAGT 2520
GCACCGATCA CAGGAGCATC GGCCGCTTCT GTGAACACTG CCCCACCTGT TATACAGCCT 2580
GCAAGGAAAA CTGGAATGTT ATGCAATGCC TTCACCTCA CAATTGTCTC CAGGCTATAC 2640
TTGATCAGTG CAAAACCTCA TGTGCTCTCA TGGAAACAAC GCATTATGTC GACCAAACTT 2700
CAGAATGTTT CTCCAGCCCA AGCTACTTGA GAATATTTT CATCATTTTC ATAGTTACAT 2760
TCTTGTATGG GTTGCTTAAA GTCTGATCA TTAGACAGGT GATACTACAA TGGAAATAGTA 2820
ATAAATATAA GTCTCATCA GATTACAGAG TGTACGCTC AAAAAAGGAT AAGTTGATTC 2880
TGCAAGATGT TTGCACAAGA GCAGTCACCT ACCGAGTGA GAAGCCTGAA GAAATAAAAA 2940
TGGATATCAG CAAATTAAT GCTCATGAAA CTTTCAGGTG CAACTTCTAA AAAAAGATT 3000
TTAAACACTT AATGGGAAAC TGGAAATGTT AATAATTGCT CCTAAGATT ATAAATTTAA 3060
AAGTCACAGG AGGAGACAAA TTGCTCACGG TCATGCCAGT TGCTGGTTGT ACATCGAAC 3120
GAAGACTGAC AAGTATCCTC ATCATGATG TACTCACATA GCTGCTGACT TTTTCAGAGA 3180
AAAATGTGTC TTAATCTGTT TTGAGACTAG TGTGCTGTA GCACCTTACT GTAATATATA 3240
ACTTATTTAG ATCAGCATAG AATGTAGATC CTCTGAAGAG CACTGATTAC ACTTTACAGG 3300
TACCTGTAT TCTCAGCTTT CCCAGAGAGA ACAATGCTGT GAGAGAGTTT AGCATTGTGT 3360
CACTACAGG GTACAGTAAT CCTTGCAGT GACATGTGAG GAAAAAATA ATCTGGCAAG 3420
TATATCTTAA GGTTCGCAAA CACTTCAACA GTTGGTGGTT GAATAGACAA GAACAGCTAG 3480
ATGAATAAAT GATTCTGTGT TCACTCTTTC AAGAGGTGAA CAGATACAAC CTTAATCTTA 3540
AAAGATTATT GCTTTTAAA GTGTGTAGTT TTATGCATGT GTGTTTATGG TTTGCTTATT 3600
TTTGCAAGAT GGATACATAA TCCAGCATTC TCTCCTCTT GCCTTTATGT TTTGTTTCT 3660
TTTTTACAGG ATAAGTTTAT GTATGTACA GATGACTGGA TTAATTAAGT GCTAAGTTAC 3720
TACTGCCATA AAAAATAAT AATACAATGT CACTTTATCA GAATACTAGT TTTAAAAGCT 3780
GAATGTTAA

Seq ID NO: 645 Protein sequence
Protein Accession #: NP_002205

45
50
55
60

1 11 21 31 41 51
MCGSALAFPT AAFVCLQND RGPASFLWAA WVFLVLGLG QGEDNRCASS NAASCARCLA 60
LGPEOCWCVQ EDFISGGRS ERCDIVSNLI SKGCSVDSIE YPSVHVLIPT ENEINTQVTP 120
GEVSIQLRPG AENFMLKVH PLKKYPVDLY YLVDVSASM NNIKLSVNG NDLSRKMAFP 180
SRDFRLGPGS YVDKTVSPYI SIHPERIHNG CSDYNLDGMP PHGYIHVLSL TENITEPEKA 240
VHRQKISGNI DTPGEGFDAM LQAAVCESHI GWRKEAKRLI LVMTDQTSHL ALDSKLALIV 300
VPNDGMCBLK NNYYVKSTM EHPSLGQLE KLIDNNINVI FAVGKQFHW YKDLLPLLP 360
TIAGEIESKA ANLNNLVVEA YQKLISEKV QVENOVQGIY FNITAI CPDG SRKPGMEGCR 420
NVTSENDELV NVTVMKKCD VTGKNYAI KPIGFNETAK IHIHRNCSQ CEDNRGPKGK 480
CVDETFLDSK CPQDENKHC FDEQFSSES CKSHKDQVPC SGRGVCVCGK CSCHKIKLKG 540
VYKGYCEKDD FSPCYHHGML CAGHGECEAG RCQCPGSGWG DRQCPGSAHA QHCVNSKGV 600
CSGRGTVCVG RCCTDPRSI GRFCEHCPTC YTACKENWNC MQCLHPNLS QAILDQCKTS 660
CALMEQHVY DQSECFSSP SYLRIFFIIF IVTFILGLLK VLIIRQVILQ WNSNKKXSS 720
DYRVSASKKD KLILQSVCTR AVTYRREKPE EIKMDISKLN AHETFRCNF

Seq ID NO: 646 DNA sequence
Nucleic Acid Accession #: NM_003318.1
Coding sequence: 1..2574

65
70
75
80
85

1 11 21 31 41 51
ATGGAATCCG AGGATTTAAG TGGCAGAGAA TTGACAATTG ATTCCATAAT GAACAAAGTG 60
AGAGACATTA AAAATAAGTT TAAAAATGAA GACCTTACTG ATGAACTAAG CTTGAATAAA 120
ATTTCTGCTG ATACTACAGA TAACTCGGGA ACTGTTAACC AAATTATGAT GATGGCAAAAC 180
AAACCAGAGG ACTGGTTGAG TTTGTTGCTC AAACAGAGAG AAAACAGTGT TCCGCTAAGT 240
GATGCTCTTT TAAATAAATT GATTGGTCTG TACAGTCAAG CAATTGAAGC GCTTCCCCCA 300
GATAAATATG GCCAAATGA GAGTTTGTCT AGAATTCAAG TGAGATTGTC TGAATTAATA 360
GCTATTCAAG AGCCAGATGA TGCACGTGAC TACTTTCAAA TGGCCAGAGC AAATGCAAG 420
AAATTTGCTT TTGTTTATAT ATCTTTTGCA CAATTGAAC TGTCAACAAG TAATGTCAAA 480
AAAAGTAAAC AACTTCTTCA AAAAGCTGTA GAACTGGAG CAGTACCCTC AGAAATGCTG 540
GAAATTTGCC TGGGGAATTT AAACCTCCAA AAAAAGCAGC TGCTTTTCAA GGAGGAAAAG 600
AAGAATTTAT CAGCATCTAC GGTATTAACT GCCCAAGAAAT CATTTTCCGG TTCATTGGG 660
CAATTACAGA ATAGGAACAA CAGTTGTGAT TCCAGAGGAC AGACTACTAA AGCCAGGTTT 720
TTATATGGAG AGAATATGCC ACCACAAGAT GCAGAAATAG GTTACCGGAA TTCATTGAGA 780
CAAACTAACA AAACATAACA GTCATGCCCA TTTGGAAGAG TCCAGTTTAA CCTTCTAAAT 840
AGCCCATAGT GTGATGTGAA GACAGATGAT TCAGTTGTAC CTTGTTTAT GAAAGACAA 900
ACCTCTAGAT CAGAAATCCG AGATTGCTT GTGCTGGAT CTAACCAAG TGGAAATGAT 960
TCCTGTGAAT TAAGAAATTT AAAGTCTGTT CAAATAGTCT ATTTCAAGGA ACCTCTGGTG 1020
TCAGATGAAA AGAGTTCTGA ACTTATTATT ACTGATTCAA TAACCTGGA GAATAAAACG 1080
GAATCAAGTC TTCTAGCTAA ATTAGAAGAA ACTAAGAGT ATCAAGAAC AGAGGTTCCA 1140
GAGAGTAACC AGAAACAGTG GCAATCTAAG AGAAAGTCAAG AGTGTATTAA CCAGAACTCT 1200
GCTGCTCTCT CAAATCACTG CGAGATTCCG GAGTTAGCCC GAAAGTTAA TACAGAGCAG 1260
AAACATACCA CTTTGTGACA ACCTGTCTTT TCAGTTTCAA AACAGTCACC ACCAATATCA 1320
ACATCTAAAT GGTTTGACCC AAAATCTATT TGTAAGACAC CAAGCAGCAA TACCTTGGAT 1380

	GATTACATGA	GCTGTTTTAG	AACGCCAGTT	GTAAGAATG	ACTTTCCACC	TGCTTGTGAG	1440
	TTGTCAACAC	CTTATGGCCA	ACCTGCCCTGT	TTCCAGCAGC	AACAGCATCA	AATACCTGCC	1500
	ACTCCACTTC	AAATTTTACA	GGTTTTAGCA	TCTTCTTCAG	CAAAATGAATG	CATTTCGGTT	1560
5	AAAGGAAGAA	TTTATTCCAT	TTTAAAGCAG	ATAGGAAGTG	GAGGTTCAAG	CAAGGTATTT	1620
	CAGGTGTTAA	ATGAAAAGAA	ACAGATATAT	GCTATAAAAT	ATGTGAACCT	AGAAGAAGCA	1680
	GATAACCAAA	CTCTTGATAG	TTACCGGAAC	GAAATAGCTT	ATTTGAATTA	ACTACAACAA	1740
	CACAGTGATA	AGATCATCCG	ACTTTATGAT	TATGAAATCA	CGGACCAAGT	CATCTACATG	1800
	GTAATGGAGT	GTGGAAATAT	TGATCTTAAT	AGTTGGCTTA	AAAAGAAAAA	ATCCATTGAT	1860
	CCATGGGAAC	GCAAGAGTTA	CTGGAATAAT	ATGTTAGAGG	CAGTTCACAC	AATCCATCAA	1920
10	CATGGCATTG	TTCCAGTGTA	TCTTAAACCA	GCTAACTTTC	TGATAGTTGA	TGGAATGCTA	1980
	AAGCTAATTG	ATTTTGGGAT	TGCAAAACCA	ATGCAACACG	ATACAACAAG	TGTTGTTAAA	2040
	GATTCACAGG	TTGGCACAGT	TAATTATATG	CCACCAGAAG	CAATCAAAGA	TATGTCTTCC	2100
	TCCAGTAGAT	ATGGCAAAATC	TAAGTCAAAG	ATAAGCCCCA	AAAGTGATGT	TTGGTCTCTA	2160
	GGATGTATT	TGTACTATAT	GACTTACGGG	AAAACACCAT	TTCAGCAGAT	AATTAATCAG	2220
15	ATTTCTAAAT	TACATGCCAT	AATTGATCCT	AATCATGAAA	TTGAATTTCC	CGATATTCCA	2280
	GAGAAAGATG	TTCAAGATGT	GTTAAAGTGT	TGTTTAAAAA	GGGACCCAAA	ACAGAGGATA	2340
	TCCATTCCCTG	AGCTCTCGGC	TCATCCCTAT	GTTCAAATTC	AAACTCATCC	AGTTAAACCA	2400
	ATGGCCAAAG	GAACCACTGA	AGAAATGAAA	TATGTTCTGG	GCCAACTGTG	TGGTCTGAAT	2460
20	TCTCCTAAAT	CCATTTTGAA	AGCTGCTAAA	ACTTTATATG	AACACTATAG	TGGTGGTGAA	2520
	AGTCATAATT	CTTCATCTTC	CAAGACTTTT	GAAAAAATAA	GGGGAAAAAA	ATGA	

Seq ID NO: 647 Protein sequence
Protein Accession #: NP_003309.1

25	1	11	21	31	41	51	
	MESEDLSGRE	LTIDSIMNKV	RDINKKFKNE	DLTDELSLNK	ISADTTDMSG	TVNQIMMMAN	60
	NPEDMLSLLL	KLEKNVPLIS	DALLNKLIGR	YSQAIEALPP	DKYQGNESFA	RIQVRFALBK	120
	AIQEPDARD	YFQMARANCK	KPAFVHISFA	QFELSQGNVK	KSKQLLQKAV	ERGAVPLEML	180
30	EIALRNLIQ	KKQLLSEEEK	KNLSASTVLT	AQESFSGSLG	HLQNRNNSCD	SRGQTTKARF	240
	LYGENMPPDG	AEIGYRNSLR	QTNKTKQSCP	PGRVVFNLLN	SPDCDVKTDD	SVVPCFMKRQ	300
	TSRSECRDLV	VPGSKPSGMD	SCELRNLSKV	QNSHFKEPLV	SDEKSSSELII	TDSITLKNKT	360
	ESSLLAKLEE	TKEYQEPFVP	ESNQKQWQSK	RKSECINQNP	AASSNHWQIP	ELARKVNTBQ	420
	KHTTFEQPVF	SVSKQSPFIS	TSKWDFPKSI	CKTPSSNTLD	DYMSCPRTPV	VKNDFPPACQ	480
35	LSTPYQQAQ	PQQQHQHILA	TPLQNLQVLA	SSSANECISV	KGRYISILKP	IGSGGSSKVF	540
	QVLNEKKQIY	AIKYVNLLEA	DNQTLDSYRN	EIAYLNKLQ	HSDKIIRLYD	YEITDQYIYM	600
	VMECGNIDLV	SWLKKKSID	PWERKSYWKN	MLEAVHTIHQ	HGIVHSDLPK	ANFLIVDGLM	660
	KLIDFGIANQ	MQPDTSVSVK	DSQVGTVNYM	PPEAIKDMSS	SRENGKSKSR	ISPKSDVWSL	720
	GCLLYMYTIG	KTFPQIINQ	ISKLHAIIDF	NHEIEFPDIP	EKDLQDVLC	CLKRDPKQRI	780
40	SIPLELAHFP	VQIQTHPVNQ	MAKGTTEEMK	YVLGLVLGLN	SPNSILKAAK	TLYEHYSGGE	840
	SHNSSSSKTF	EKKRGKK					

Seq ID NO: 648 DNA sequence
Nucleic Acid Accession #: NM_015507
Coding sequence: 241..1902

45	1	11	21	31	41	51	
	CGCAGAGGA	GCCTCGGCCA	GGCTAGCCAG	GGCGCCCCCA	GCCCCCTCCC	AGGCCGOGAG	60
	CGCCCCGCGC	GGCGTGCCCTG	GCCTCCCCCTG	CCAGACTGCA	GGGACAGCAC	CCGGTAACTG	120
	CGAGTGAGAG	GGAGGACCCG	AGCGGCTGAG	GAGAGAGGAG	GCGGCGGCTT	AGCTGCTAAG	180
	GGGTCCGGCC	GCGCCCTCTG	CGAGGGGGGC	TCAGGAGGAG	GAAGGAGGAC	CCGTGCGAGA	240
	ATGCTCTGCG	CCTGGAGCCT	TGCGCTCCCG	CTGCTGCTCT	CCTGGGTGGC	AGGTGGTTTC	300
	GGGAACGGGG	CCAGTGCAAG	GCATCACGGG	TGTTTAGCAT	CGGCACGCTA	GCCTGGGGTC	360
55	TGTCACTATG	GAACTAACTG	GGCCTGCTGC	TACGGCTGGA	GAAGAAACAG	CAAGGGAGTC	420
	TGTGAAGCTA	CATGCGAAC	TGGATGTAAG	TTTGGTGAGT	GCGTGGGACC	AAACAAATGC	480
	AGATGCTTTC	CAGGATACAC	CGGAAAAACC	TGCAGTCAAG	ATGTGAATGA	GTGTGGAAATG	540
	AAACCCCGCG	CATGCCAACCA	CAGATGTGTG	AATACACACG	GAAGCTACAA	GTGCTTTTGC	600
	CTCAGTGGCC	ACATGCTCAT	GCCAGATGCT	ACGTGTGTGA	ACTCTAGGAC	ATGTGCCATG	660
60	ATAAACTGTC	AGTACAGCTG	TGAAGACACA	GAAGAAGGGC	CACAGTGCCT	GTGTCCATCC	720
	TCAGGACTCC	GCTCGCCCCC	AAATGGAAGA	GACTGTCTAG	ATATTGATGA	ATGTGCCTCT	780
	GGTAAAGTCA	TCTGTCCCTA	CAATCGAAGA	TGTGTGAACA	CATTGTGAAG	CTACTACTGC	840
	AAATGTGACA	TTGTTTTCGA	ACTGCAATAT	ATCAGTGGAC	GATATGACTG	TATAGATATA	900
	AATGAATGTA	CTATGATATG	CCATACGTGC	AGCCACCATG	CCAATTGCTT	CAATACCCAA	960
65	GGGTCTCTCA	AGTGTAAATG	CAAGCAGGGA	TATAAAGGCA	ATGGACTTCG	GTGTTCTGCT	1020
	ATCCCTGAAA	ATTCTGTGAA	GGAAGTCCTC	AGAGCACCTG	GTACCATCAA	AGACAGAAATC	1080
	AAGAAGTTGC	TGTCTCACAA	AAACAGCATG	AAAAAGAAGG	CAAAAATTAA	AAATGTTACC	1140
	CCAGAACCCA	CCAGGACTCC	TACCCCTAAG	GTGAACCTGC	AGCCCTTCAA	CTATGAAGAG	1200
	ATAGTTTCCA	GAGGCGGGAA	CTCTCATGGA	GGTAAAAAAG	GGAATGAAGA	GAAATGAAAG	1260
70	GAGGGGCTTG	AGGATGAGAA	AAGAGAAGAG	AAAGCCCTGA	AGAATGACAT	AGAGGAGCGA	1320
	AGCCTGCGAG	GAGATGTGTT	TTTCCCTAAG	GTGAATGAAG	CAGGTGAATT	CGGCCTGATT	1380
	CTGGTCCAAA	GGAAAGCGCT	AACCTCCAAA	CTGGAACATA	AAGATTTAAA	TATCTCGGTT	1440
	GACTGCGAGT	TCAATCATGG	GATCTGTGAC	TGGAAACAGG	ATAGAGAAGA	TGATTTTGAC	1500
	TGGAATCCCTG	CTGATCGAGA	TAATGCTATT	GGCTTCTATA	TGGCAGTTCC	GGCCTTGGCA	1560
75	GGTCACAAGA	AAGACATTGG	CCGATTGAAA	CTTCTCTTAC	CTGACCTGCA	ACCCCAAAGC	1620
	AACCTCTGTT	TGCTCTTTGA	TTACCGGCTG	GCCGGAGACA	AAGTCGGGAA	ACTTCGAGTG	1680
	TTTGTGAAAA	ACAGTAACAA	TGCCCTGGCA	TGGGAGAAGA	CCAAGAGTGA	GGATGAAAAG	1740
	TGGAAGACAG	GGAAAATTCA	GTTGTATCAA	GGAACTGATG	CTACCAAAAG	CATCATTTTT	1800
	GAAGCAGAAC	GTGGCAAGGG	CAAAACCGGC	GAAATCGCAG	TGGATGGCGT	CTTGCTTGTT	1860
80	TCAGGCTTAT	GTCCAGATAG	CTTTTATCT	GTGGATGACT	GAAATGTTACT	ATCTTTATAT	1920
	TTGACTTTGT	ATGTCAGTTC	CCTGGTTTTT	TGATATTGTC	ATCATAGGAC	CTCTGGCATT	1980
	TTAGAATTAC	TAGCTGAAAA	ATTGTAATGT	ACCAACAGAA	ATATTATTGT	AAGATGCGCT	2040
	TCTTGTATAA	GATATGCCAA	TATTTGCTTT	AAATATCATA	TCAGTGTATC	TTCTCAGTCA	2100
	TTTCTGAATC	TTTCCACTAT	ATATTATAAA	ATATGGAAT	GTGAGTTTAT	CTCCCTCTCT	2160
85	CAGTATATCT	GATTGTATA	AGTAAGTTGA	TGAGCTTCTC	TCTACAACAT	TTCTAGAAAA	2220
	TAGAAAAAAA	AGCACAGAGA	AATGTTTAAC	TGTTTGACTC	TTATGATACT	TCTTGGAAAC	2280
	TATGACATCA	AAGATAGACT	TTTGCCTAAG	TGGCTTAGCT	GGGTCTTTCA	TAGCCAAACT	2340

TGTATATTTA AATCTCTTGT AATAATAATA TCCAAATCAT CAAAAAATAA AAAAAAAA

Seq ID NO: 649 Protein sequence
Protein Accession #: NP_056322

5

1	11	21	31	41	51	
MPLPWSLALP	LLLSWVAGGF	GNAASARHHG	LLASARQPGV	CHYGTKLACC	YGWRRNSKGV	60
CEATCEPGCK	FGSCVGNPKC	RCPPGYTGKT	CSQDVNECGM	KPRPCQHRVC	NTHGSYKCFK	120
LSGHNLMPPDA	TCVNSRTCAM	INQYSCEDT	EEGPQCLCPK	SGLRLAPNGR	DCLDIDECAS	180
GKVICPNRNR	CVNTFGSYYC	KCHIGFELQY	ISGRYDCIDI	NECTMDSHTC	SHHANCPTNQ	240
GSFKCKCKQG	YKGNGLRCSA	IPENSVKEVL	RAPGTIKDRI	KGLLAHKNMS	KQKAKIKNVT	300
PEPTRTPFPK	VNLQPNVEEB	IVSRGGNSHG	GKKGNEEFMK	EGLEDEKREE	KALKNDIEER	360
SLRGDVFFPK	VNEAGEFGLI	LVQRKALTSK	LEHKDLNISV	DCSPNHGICD	WKQDREDDFD	420
WNPADRDNAI	GFYMAVPALA	GHKKDIGRLK	LLLPDLQPQS	NFCILLFDYRL	AGDKVKGKLRV	480
FVKNSNNALA	WKTTSSEDEK	WKTGKIQLYQ	GTDATKSIIF	EAERGGKGTG	EIAVDGVLLV	540
SGLCPSLSLS	VDD					

20

Seq ID NO: 650 DNA sequence
Nucleic Acid Accession #: NM_003506.1
Coding sequence: 259..2379

25

1	11	21	31	41	51	
GCAGCTCCAG	TCCCGGACGC	AACCCCGGAG	CCGTCTCAGG	TCCCTGGGGG	GAACCGTGGG	60
TTAGAGGGGG	ACGGGAAGGG	ACAGCGGCCT	TCCACCGCCC	CCCGAGTAAT	TGACCCAGGA	120
CTCATTTTCA	GGAAGCCTGT	AAAATGAGTA	AAATAGTGAA	ATGAGGAATT	TGAACATTTT	180
ATCTTTGGAT	GGGGATCTTC	TGAGGATGCA	AAGAGTGATT	CATCCAAGCC	ATGTGTGAAA	240
ATCAGGAATT	TGAGAGAAAT	GGAGATGTTT	ACATTTTGTG	TGACGTGTAT	TTTCTACCC	300
CTCCTAAGAG	GGCAGACGCT	CTTCACCTGT	GAACCAATTA	CTGTTCCCGA	ATGTATGAAA	360
ATGGCCTACA	ACATGACGTT	TTTCCCTAAT	CTGATGGGTC	ATTATGACCA	GAGTATTGCC	420
GCGGTGGAAA	TGAGACATTT	TCCTCCTCTC	GCAAACTCTG	AATGTTCAAC	AAACATTGAA	480
ACTTTCCTCT	GCAAGCATT	TGTACCAACC	TGCATAGAAC	AAATTCATGT	GGTTCACCT	540
TGTCGTAAAC	TTGTGTGAGAA	AGTATATTCT	GATTGCAAAA	AATTAATTGA	CACITTTGGG	600
ATCCGATGGC	CTGAGGAGCT	TGAATGTGAC	AGATTACAAT	ACTGTGATGA	GACTGTTCTT	660
GTAACTTTTG	ATCCACACAC	AGAATTCTCT	GGTCTCTAGA	AGAAAACAGA	ACAAGTCCAA	720
AGAGACATTG	GATTTTGGTG	TCCAAGGCAT	CTTAAGACTT	CTGGGGGACA	AGGATATAAG	780
TTTCTGGGAA	TTGACCAGTG	TGCGCCTCCA	TGCCCCAACA	TGTATTTTAA	AAGTGATGAG	840
CTAGAGTTTG	CAAAAAGTTT	TATTGGAACA	GTTTCAATAT	TTTGTCTTTG	TGCAACTCTG	900
TTACATTTCC	TTACTTTTTT	AATTGATGTT	AGAAGATTCA	GATACCCAGA	GAGACCAATT	960
ATATATTACT	CTGTCTGTGA	CAGCAATTGA	TCTCTTATGT	ACTTCAATTG	ATTTTGTGCT	1020
GGCGATAGCA	CAGCCTGCAG	TAAGGCAGAT	GAGAAGCTAG	AACCTGTGTA	CACTGTTGTC	1080
CTAGGCTCTC	AAAATAAGGC	TTGCACCGTT	TTGTTCATGC	TTTTGTATTT	TTTCACAAAT	1140
GCTGGCACTG	TGTGGTGGGT	GATTCCTACC	ATTACTGGGT	TCTTAGCTGC	AGGAAGAAAA	1200
TGGAGTTTGG	AAGCCATCGA	GCAAAAAGCA	GTGTGGTTTC	ATGCTGTGTC	ATGGGGGAACA	1260
CCAGGTTTCC	TGACTGTTAT	GCTTCTTGCT	CTGAACAAAG	TTGAAGGAGA	CAACATTAGT	1320
GGAGTTTGCT	TTGTTGGCCT	TTATGACCTG	GATGCTTCTC	GCTACTTTGT	ACTCTTGCCA	1380
CTGTGCTTTT	GTGTGTTTGT	TGGGCTCTCT	CTTCTTTTAG	CTGGCATTAT	TTCTTTAAAT	1440
CATGTTTCAG	AAGTCATACA	ACATGATGGC	CGGAACCAAG	AAAAACTAAA	GAATTTTATG	1500
ATTGCAATTG	GAGTCTTCAG	CGGCTGTGAT	CTTGTGCCAT	TAGTGACACT	TCTCGGATGT	1560
TACGTCATGT	AGCAAGTGAA	CAGGATTACC	TGGGAGATAA	CTTGGGTCTC	TGATCATGTG	1620
CGTCAGTACC	ATATCCCATG	TCCTTATCAG	GCAAAAGCAA	AAGCTCGACC	AGAATTGGCT	1680
TTATTTATGA	TAAATACCTT	GATGACATTA	ATTGTTGGCA	TCTCTGCTGT	CTTCTGGGTT	1740
GGAGGCAAAA	AGACATGCAC	AGAAATGGGCT	GGGTTTTTTA	AACGAAATCG	CAAGAGAGAT	1800
CCAATCAGTG	AAAGTCGAAG	AGTACTACAG	GAATCATGTG	AGTTTTTCTT	AAAGCACAAT	1860
TCTAAAGTTA	AACACAAAAA	GAAGCACTAT	AAACCAAGTT	CACACAAGCT	GAAGGTCAAT	1920
TCCAAATCCA	TGGGAACCCG	CACAGGAGCT	ACAGCAATTC	ATGGCACTTC	TGCAGTAGCA	1980
ATTACTAGCC	ATGATTACCT	AGGACAGAA	ACTTTGACAG	AAATCCAAAC	CTCACCAGAA	2040
ACATCAATGA	GAGAGGTGAA	AGCGGACGGA	GCTAGCACCC	CCAGGTTAAG	AGAACAGGAC	2100
TGTGTGAAC	CTGCTCGGCC	AGCAGCATCC	ATCTCCAGAC	TCTCTGGGGA	ACAGGTGAGC	2160
GGGAAGGGCC	AGGCAGGCAG	TGTATCTGAA	AGTGCAGGGA	GTGAAGGAAG	GATTAGTCCA	2220
AAGAGTGATA	TTACTGACAC	TGGCCTGGCA	CAGAGCAACA	ATTTCAGGTT	CCCCAGTTCT	2280
TCAGAACCAA	GCAGCCTCAA	AGGTTCCACA	TCTCTGCTTG	TTCACCCACT	TTCAAGGAGT	2340
AGAAAAGAGC	AGGAGGTTGG	TTGTCAATTA	GATACCTTGA	GAACATTTTC	TCTCGTTACT	2400
CAGAAGCAAA	TTTGTGTTAC	ACTGGAAGTG	ACCTATGCAC	TGTTTTGTAA	GAATCACTGT	2460
TACGTTCTTC	TTTTGCACTT	AAAGTGCAT	TGCTACTGT	TATACTGGAA	AAAAATAGAT	2520
TCAGAGATAA	TATGACTCAT	TTCAACAAAA	GGTTAATGAC	AACAATATAC	CTGAAAAACG	2580
AAATGTGCAG	GTTAATAATA	TTTTTTTAAT	AGTGTGGGAG	GACAGAGTTA	GAGGAATCTT	2640
CTTTTCTAT	TTATGAAGAT	TCTACTCTTG	GTAAGAGTAT	TTTAAGATGT	ACTATGCTAT	2700
TTTACCTTTT	TGATATAAAA	TCAAGATATT	TCTTTGCTGA	AGTATTTAAA	TCTTATCCTT	2760
GTATCTTTTT	ATACATATT	GAAAAAAGC	TTATATGTAT	TTGAACTTTT	TTGAAATCCT	2820
ATTCAAGTAT	TTTTATCATG	CTATTGTGAT	ATTTTAGCAC	TTTGTAGCT	TTTACACTGA	2880
ATTTCTAAGA	AAATTGTAAA	ATAGTCTTCT	TTTATACTGT	AAAAAAGAT	ATACCAAAAA	2940
GTCTTATAAT	AGGAATTTAA	CTTTAAAAAC	CCACTTATTG	ATACCTTACC	ATCTAAAAATG	3000
TGTGATTTT	ATAGTCTCGT	TTTAGGAATT	TCACAGATCT	AAATTATGTA	ACTGAAATAA	3060
GGTGCTTACT	CAAGAGGTGT	CCACTATTGA	TTGTATTATG	CTGCTCACTG	ATCCTTCTGC	3120
ATATTTAAAA	TAAAAATGTC	TAAAGGGTTA	GTAGACAAAA	TGTTAGTCTT	TTGTATATTA	3180
GGCCAGTGCC	AAATTGACTC	CCTTTTTTAA	TGTTTATGTA	CCACCCATG	ATTGTATTAT	3240
AACCACTTAC	AGTTGCTTAT	ATTTTTTGT	TTAACTTTTG	TTTCTTAAAC	TTTGAATAT	3300
TACATTTTGT	ATTATACAGT	ACCTTTCTCA	GACATTTTGT	AG		

85

Seq ID NO: 651 Protein sequence
Protein Accession #: NP_003497.1

1	11	21	31	41	51

MEMFTLLTLC IFPLLRGHS LFTCEPITVP RCMKMYNMT FFPNLMGHYD QSIAAEMEH 60
 FLPLANLECS PNIEITFLCKA FVPTCEIQIH VVPPCRKLCE KVSDDCKLI DTFGIRNPRE 120
 LECDRLYQCD ETVPVTFDPH TEFLGPQKKT EQVQRDIPGW CPRHLKTSYG QGYKFLGIDQ 180
 CAPPCFNMYF KSDELEFAKS FIGTVSIFCL CATLFTFLTP LIDVRRFRYP ERPIIYSVC 240
 5 YSIVSLWYFI GFLLGDSSTAC NKADEKLELQ DTVVLGSQNK ACTVLFMLLY PFTMAGTVWN 300
 VILITITWFLA AGRKMSCEAI EQKAVWFHAV AWGTPGFLLTV MLLALNKVEG DNISGVCFVG 360
 LYDLASRYF VLLPLCLCVF VGLSLLLAGI ISLNVHRQVI QHDGRNQEKI KCPMIRIGVP 420
 SGLYLVPLVT LLGCVYVEQV NRITWEITWV SDHCRQYHIP CPYQAKAKAR PELALPMIKY 480
 10 LMTLIUVISA VFWVGSKKTC TEWAGFFKRN RKRDPISERS RVLQESCEFP LKHSKVKHK 540
 KKHYPSSHK LKVISKSMGT STGATANHGT SAVAITSHDY LGQETLLEIQ TSPETSMREV 600
 KADGASTPRL REQDCGEPAS PAASISRLSG EQVDGKGQAG SVSESARSEG RISPKSDITD 660
 TGLAQSNLQ VPSSSEPSL KGSTSLLVHP VSGVRKEQGS GCHSDT

Seq ID NO: 652 DNA sequence
 Nucleic Acid Accession #: NM_014791.1
 Coding sequence: 171..2126

1 11 21 31 41 51
 20 TTGGCGGGCG GAAGCGGCCA CAACCCGGCG ATCGAAAAGA TTCTTAGGAA CGCGGTACCA 60
 GCCGCTCTC TCAGACACAG AGGCCCTCTG CCTTCTGTG GCGCGCGCTC AGCCGTGCCC 120
 TCGGCCCTC AGGTCTCTTT TCTAATTCCA AATAAACTTG CAAGAGGACT ATGAAAGATT 180
 ATGATGAATC TCTCAAATAT TATGAATTAC ATGAAACTAT TGGGACAGGT GGCCTTGCAA 240
 25 AGGTCAAATC TGCTGCCAT ATCCTTACTG GAGAGATGGT AGCTATAAAA ATCATGGATA 300
 AAAACACACT CAGGAGTGAT TTGCCCCGGA TCAAAAAGGA GATTGAGGCC TTGAAGAACC 360
 TGAGACATCA GCATATATGT CAACTCTACC ATGTGCTAGA GACAGCCAAC AAAATATTCA 420
 TGGTTCTTGA GTCTGCTCTT GGAGGAGAGC TGTTTGACTA TATAATTCC CAGGATCGCC 480
 TGTGAGAAGA GSAGACCCGG GTTGTCTTCC GTCAGATAGT ATCTGCTGTT GCTTATGTGC 540
 30 ACAGCCAGGG CTATGCTCAC AGGGACCTCA AGCCAGAAAA TTTGCTGTTT GATGAATATC 600
 ATAAATTAAG GCTGATTGAC TTTGGTCTCT GTGCAAAACC CAAGGGTAAC AAGGATTACC 660
 ATCTACAGAC ATGCTGTGGG AGTCTGGCTT ATGCAGCACC TGAGTTAATA CAAGGCAAAAT 720
 CATATCTGGG ATCAGAGGCA GATGTTTGA GCATGGGCAT ACTGTTATAT GTTCTTATGT 780
 GTGGATTCTC ACCATTTGAT GATGATAATG TAATGGCTTT ATACAAGAAG ATTATGAGAG 840
 35 GAAAATATGA TGTTCCTCAAG TGGCTCTCTC CCGATAGCAT TCTGCTTCTT CAACAAATGC 900
 TGCAGGTGGA CCCAAAGAAA CGGATTCTTA TGAAAAATCT ATTGAACCAT CCCTGGATCA 960
 TGCAGATTGA CAATATCTCT GTTGAGTGGC AAAGCAAGAA TCCTTTTATT CACCTCGATG 1020
 ATGATTGGCT AACAGAACTT TCTGTACATC ACAGAAACAA CAGGCAAAAC ATGGAGGATT 1080
 TAATTTCACT GTGGCAGTAT GATCACCTCA CGGCTACCTA TCTTCTGCTT CTAGCCAAGA 1140
 40 AGGCTCGGGG AAAACACAGT CGTTTAAGGC TTTCTTCTTT CTCTGTGGA CAAGCCAGTG 1200
 CTACCCCATC CACAGACATC AAGTCAAATA ATTGGAGTCT GGAAGATGTG ACCGCAAGTG 1260
 ATAAAAATTA TGTGGCGGGA TTAATAGACT ATGATTGGTG TGAAGATGAT TTATCAACAG 1320
 GTGCTGCTAC TCCCGGAACA TCACAGTTTA CCAAGTACTG GACAGAATCA AATGGGGTGG 1380
 AATCTAAATC ATTAACCTCA GCCTTATGCA GAACACCTGC AAATAAATTA AAGAACAAAG 1440
 45 AAAATGTATA TACTCTTAAG TCTGCTGTAA AGAATGAAGA GTACTTTATG TTTCTGAGC 1500
 CAAAGCTCCG AGTTAATAAG AACCAAGCATA AGAGAGAAAT ACTCACTACG CCAAAATGTT 1560
 ACACTACACC CTCAAAGAGT AGAAACAGT GCCTGAAAGA AACTCCAATT AAAATACCAG 1620
 TAAATCAAC ASGAACAGAC AAGTTAATGA CAGGTGTCAT TAGCCCTGAG AGCGGGTGCC 1680
 GCTCAGTGGG ATTGGATCTC AACCAAGCAC ATATGGAGGA GACTCCAAAA AGAAAGGGAG 1740
 50 CCAAGTGTGT TGGGAGCCTT GAAAGGGGGT TGGATAAGGT TATCACTGTG CTCACCAAGA 1800
 GCAAAAGGAA GGGTCTGCGC AGAGACGGGC CCAGAAGACT AAAGCTTCAC TATAATGTGA 1860
 CTACAACTAG ATTAGTGAAT CCAGATCAAC TGTGAATGA AATAATGTCT ATTCTTCAA 1920
 AGAAGCATGT TGACTTTGTA CAAAAGGGTT ATACACTGAA GTGTCAAACA CAGTCAGATT 1980
 TTGGGAAAGT GACAAATGCA TTTGAATTAG AAGTGTGCCA GCTTCAAAAA CCGGATGTGG 2040
 TGGGTATCAG GAGGACGGCG CTTAAGGGCG ATGCCTGGGT TTACAAAAGA TTAGTGGAAG 2100
 55 ACATCTATC TAGCTGCAAG GTATAATTGA TGGATTCTTC CATCTGCCG GATGAGTGTG 2160
 GGTGTGATAC AGCTACATA AAGACTGTGA TGATCGCTTT GATTTTAAAG TTCATTGGAA 2220
 CTACCAACTT GTTCTTAAG AGCTATCTTA AGACCAATAT CTCTTTGTTT TTAACAAAA 2280
 GATATTATTG TGTGATGAA TCTAAATCAA GCCATCTGT CATTATGTGA CTGCTTTT 2340
 60 TAATCATGTG GTTTGTATA TTAATAATTG TTGACTTCT TAGATTCACT TCCATATGTG 2400
 AATGTAAGCT CTTAACTATG TCTCTTTGTA ATGTGTAATT TCTTTCTGAA ATAAACCAT 2460
 TTGTGAATAT

Seq ID NO: 653 Protein sequence
 Protein Accession #: NP_055606.1

1 11 21 31 41 51
 65 MKDYDELLKY YELHETIGTG GFAKVKLACH ILTGEMVAIK IMDKNTLGSD LPRIKTEIEA 60
 LKNLRHQHIC QLYHVLETAN KIPMVLEYCP GGELFDYIIS QDRLESEETR VVFRQIVSAV 120
 70 AYVHSQGYAH RDLKPNELLF DEYHKLKID FGLCAKPKGN KDYHLQTCG SLAYAAPELI 180
 QGKSYLGSEA DVNSMGILLY VLMCGFLFPD DDNVMAKYK IMRGKYDVPK WLSPPSILL 240
 QQMLQVDPK RISMKNLLNH PWMQDYNYP VEWQSKNPF I HLDGDCVTEL SVHERNNRQT 300
 MEDLISLVQY DHLTATYLLL LAKKARGKPV RLRLSSPFCG QASATPPTDI KSNWNSLEDV 360
 75 TASDKNYVPK LIDYDWCEDD LSTGAATPRT SQFTKYWTES NGVESKSLTP ALCRTPANKL 420
 KNKENYTPK SAVKNEEYFM FPEKTPVNVK NQHKREILT I PMRYTTPSKA RNQCLKETPI 480
 KIPVNSTGTD KLMTGVISPE RRCRSVELDL NQAHMEETPK RKGAKVPSL ERGLDKVITV 540
 LTRSKRKGA RDGPRRLKIH YNVTTLVLVN PDQLNEIMS ILPKKHVPV QKGYTLKQCT 600
 QSDFGVMTQ FELEVCLQK PDVVGIRRRQ LKGDWVYKR LVEDILSSCK V

Seq ID NO: 654 DNA sequence
 Nucleic Acid Accession #: NM_000582
 Coding sequence: 88..990

1 11 21 31 41 51
 85 GCAGAGCACA GCATCGTCGG GACCAGACTC GTCTCAGGCC AGTTGCAGCC TTCTCAGCCA 60
 AACGCCGACC AAGGAAAAC TACTACCATG AGAATTGCAG TGATTGTGCT TTGCTCTCTA 120

5 GGCATCACCT GTGCCATACC AGTTAAACAG GCTGATTCTG GAAGTTCTGA GGAAGAGCAG 180
 CTTTACACCA AATACCCAGA TGCTGTGGCC ACATGGCTAA ACCCTGACCC ATCTCAGAAG 240
 CAGAATCTCC TAGCCCCACA GACCCCTCCA AGTAAGTCCA ACGAAAGCCA TGACCACATG 300
 GATGATATGG ATGATGAAGA TGATGATGAC CATGTGGACA GCCAGGACTC CATTGACTCG 360
 AACGACTCTG ATGATGTAGA TGACACTGAT GATTCTCACC AGTCTGATGA GTCTCACCAT 420
 TCTGATGAAT CTGATGAAC TGGTCACTGAT TTTCCACCGG ACCTGCCAGC AACCGAAGTT 480
 TTCACTCCAG TTGTCCCCAC AGTAGACACA TATGATGGCC GAGGTGATAG TGTGTTTAT 540
 GGACTGAGGT CAAAATCTAA GAAGTTTCOG AGACCTGACA TCCAGTACCC TGATGTCTACA 600
 10 GACGAGGACA TCACCTCACA CATGGAAGC GAGGAGTTGA ATGGTGCTA CAAGGCCATC 660
 CCGGTGGCC AGGACCTGAA CGCGCTTCT GATTGGGACA GCCGTGGGAA GGACAGTTAT 720
 GAAACGAGTC AGCTGGATGA CCAGAGTGCT GAAACCCACA GCCACAAGCA GTCCAGATTA 780
 TATAAGCGGA AAGCCCAATGA TGAGAGCAAT GAGCATTCOG ATGTGATTGA TAGTCAGGAA 840
 CTTTCCAAAG TCAGCCGTGA ATTCCACAGC CATGAATTC ACAGCCATGA AGATATGCTG 900
 GTTGTAGACC CCAAAAGTAA GGAAGAAGAT AAACACCTGA AATTCGTAT TTCTCATGAA 960
 15 TTAGATAGTG CATCTTCTGA GGTCAATTAA AAGGAGAAAA AATACAATTT CTCACITTGC 1020
 ATTTAGTCAA AAGAAAAAAT GCTTTATAGC AAAATGAAAG AGAACATGAA ATGCTTCTTT 1080
 CTGAGTTTAT TGGTTGAATG TGTATCTATT TGAGTCTGGA AATAACTAAT GTGTTTGATA 1140
 ATTAGTTTAG TTTTGGCTT CATGGAACCT CCCTGTAAC TAAAGCTTC AGGGTTATGT 1200
 20 CTATGTTTAT TCTATAGAAG AAATGCAAA TATCACTGTA TTTAATATT TGTATTCTC 1260
 TCAATGAATG AATTTATGT AGAAGCAAA AAAATACTTT TACCCACTTA AAAAGAGAA 1320
 ATAACATTTT ATGTCACTAT AATCTTTTGT TTTTAAAGTT AGTGATATAT TGTGTTGAT 1380
 TATCTTTTGT TGGTGTGAAT AAATCTTTTA TCTTGAATGT AATAAGAAAT TGGTGGTGTC 1440
 AATTGCTTAT TGTTTTCCC ACGTTGTCC AGCAATTAAT AAAACATAAC CTTTTTACT 1500
 25 GCCTAAAAAA AAAAAAAA AAAA

Seq ID NO: 655 Protein sequence

Protein Accession #: NP_000573

30 1 11 21 31 41 51
 | | | | |
 MRIAVICFLD LGITCAIPVK QADSGSSEK QLYNKYPDAV ATWLNPDPSQ KQNLAPQTL 60
 PSKSNESHDL MDMDDDEDD DHVDSQDSID SNDSDDVDDT DSHQSDSHS HSDSEDELVT 120
 DFTDLPLATE VFTPVVFTVD TYDGRGDSVV YGLRSKSKFP RRPDIQYFDA TDEDITSHME 180
 35 SEELNGAYKA IPVAQDLNAP SDWDSRGKDS YETSQLDDQS AETHSHKQSR LYKRKANDES 240
 NEHSDVIDSQ ELKSVREPH SHBFHSHEDM LVVDPKSKEE DKHLKFRISH ELDSASSEVN

Seq ID NO: 656 DNA sequence

Nucleic Acid Accession #: NM_003108.1

Coding sequence: 76..1401

40 1 11 21 31 41 51
 | | | | |
 45 GGGGTGGGAG GGGGAGGGGG ACCTCCGCAC GAGACCCAGC GGCCCGGGTT GGAGCGTCCA 60
 GCCCTGCAAC GGATCATGCT GCAGCAGGCG GAGAGCTTGG AAGCGGAGAG CAACCTGCCC 120
 CGGAGGGGCGC TGGACACGGA GGAGGGCGAA TTCATGGCTT GCAGCCCGGT GCCCTGGGAC 180
 GAGAGCGACC CAGACTGGTG CAAGACGCGC TCGGCCACA TCAAGCGGCC GATGAACGGG 240
 TTCATGGTAT GGTCCAAGAT CGAACGCAAG AAGATCATGG AGCAGTCTCC GGACATGCAC 300
 AACCGCGAGA TCTCCAAGAG GCTGGGCAAG CGCTGGAAGA TGCTGAAGGA CAGCGAGAAG 360
 50 ATCCCGTTCA TCCGGGAGCG GGAGCGGCTG CGGCTCAAGC ACATGGCCGA CTACCCCGAC 420
 TACAAGTACC GGCCCGGGA AAAGCCCAA ATGGACCCCT CGGCCAAGCC CAGCGCCAGC 480
 CAGAGCCCAAG AGAAGAGCGC GGCCGCGCGC GGCCGCGGGA GCGCGGGCGG AGCGCGGGGC 540
 GGTGCCAAGA CCTCCAAGGG CTCCAGCAAG AAATGCGGCA AGCTCAAGGC CCCCAGCGCC 600
 GCGGGCGCCA AGCGGGGCGC GGGCAAGCGC GCCCAGTCCG GGGACTACGG GGGCGGGGC 660
 55 GACGACTAGC TGTGGGCGAG CCTGCGCTGG AGCGGCTGG GCGCGGCGG GCGGGCAAG 720
 ACGTCAAGT GCGTGTCTT GATGAGGAC GACGACGAG ACGACGACGA CGACGAGCTG 780
 CAGCTGCAGA TCAACAGGA GCCCGACGAG GAGGACGAG AACCAACGCA CCACGAGCTC 840
 CTGCAGCCGC CGGGGCGCAG GCCGTGCGAG CTGCTGAGAC GCTACAACTG CGCCAAAGTG 900
 CCGCCAGGCC CTACGCTGAG CAGCTCGCGC GAGTCCCGCG AGGGAGGAGG CCTCTAGAC 960
 60 GAGGTGCGGG CGCGCGCGAG CTGCGGCGCC GGGGGCGGCA GCGGCTCTA CTACAGCTTC 1020
 AAGAACATCA CCAAGCAGCA CCGCGCGCGC CTCGCGCAGC CGCGCTGTC GCCCGCTCC 1080
 TCGCGCTCGG TGTCCACCTC CTCGTCCAGC AGCAGCGGCA GCAGCAGCG CAGCAGCGG 1140
 GAGGACGCGC ACGACCTGAT GTTGGACCTG AGCTTGAATT TCTCTCAAG CGCGCACAGC 1200
 GCGAGCGAGC AGCAGCTGGG GGGCGGCGCG GCGGCGGGA ACCTGTCCCT GTCGCTGGT 1260
 65 GATAAGGATT TGGATTCTGT CAGCGAGGGC AGCCTGGGCT CCCACTTCA GTTCCCGAC 1320
 TACTGCACGC CGGAGCTGAG CGAGATGATC GCGGGGACT GGTGAGGC GAACCTCTCC 1380
 GACCTGGTGT TCACATATTG AAAGCGCGCC GCTGCTCGCT CTTTCTCTCG GAGGGTGCAG 1440
 AGCTGGGTTT CTTGGGAGGA AGTTGTAGTG GTGATGATGA TGATGATGAT AATGATGAT 1500
 70 ATGATGGTGG TGTGTATGTT GCGGTGTTA GGGTGGAGGG GAGAGAAGAA GATGCTGATG 1560
 ATATTGATAA GATGTCGTGA CGCAAGAAA TTGGAAGAAA TGATGAAAT TTTGGTGGAG 1620
 TTAAGTGAA ATGATGATGT TTTAAACATT TTCTCTGTC TTTTCTGTC CCCCCTCCCT 1680
 TCCTTTATCG TGCTCAAGG TAGTTGCATA CCTAGTCTGG AGTTGTGATT ATTTTCCCAA 1740
 AAAATGTGTT TTTGTAATTA CTATTTCTTT TTCTGAAAT TCGTGATTGC AACAAAGGCA 1800
 75 GAGGGGCGCG CGCGCGGAG GGGAGGTAGG ACCCGCTCCG GAAGGCGCTG TTTGAAGCTT 1860
 GTCGGTCTTT GAAGCTTGA AGACGCTGCG AGAGGACCCCT TTTGGCAGCA CAACTGTTAC 1920
 TCTAGGAGAT TGGTGGAGAT ATTTTCTTTT CTTAAGAGAA CTTAAGAAC TGGTGATTTT 1980
 TTTTAAACAA AAAAGGGG

Seq ID NO: 657 Protein sequence

Protein Accession #: NP_003099.1

80 1 11 21 31 41 51
 | | | | |
 85 MVQAESLEA ESNLPREALD TEEGEFMACS FVALDESDDP WCKTASGHK RPMNAPMVWS 60
 KIERRKIMQ SPDMNAEIS KRLGKRWKML KDSEKIPFIR EAERLRKHM ADYDPDYKRP 120
 RKKPKMDPSA KPSASQSPK SAAGGGGSA GGGAGGAKTS KGSKKCKGL KAPAAAGAKA 180
 GAGKAAQSGD YGAGDDYVL GSLRVSGSGG GGAGKTVKCV FLDDEDDDDD DDELQLQIR 240

QEPDEDEEP PHQQLLQPPG QQPSQLLRRY NVAKVPASPT LSSSAESPEG ASLYDEVVRAG 300
 ATSGAGGSSR LYYSPKNITK QHPPPLAQPA LSPASSREVS TSSSSSSGSS SGSSGEDADD 360
 LMFDLSLNFPS QSAHSASBQQ LGGGAAAGNL SLSLVDKDLDF SFSEGLSGSH FEFPDYCTPE 420
 LSEMIAGDWL EANFSDLVFT Y

Seq ID NO: 658 DNA sequence
 Nucleic Acid Accession #: NM_001719
 Coding sequence: 123..1418

10 1 11 21 31 41 51
 GGGCGCAGCG GGGCCCGTCT GCAGCAAGTG ACCGACGGCC GGGACGGCCG CCTGCCCCCT 60
 CTGCCACCTG GGGCGGTGCG GGGCCGGAGC CCGGAGCCCG GGTAGCGCGT AGAGCCGGCG 120
 CGATGCACTG GCGCTCACTG CGAGCTGCGG CGCGGCACAG CTTCTGGCGG CTCTGGGCAC 180
 15 CCTGTTTCTT GCTGCGCTCC GCCCTGGCCG ACTTCAGCCT GGACAACGAG GTGCACTCGA 240
 GCTTCATCCA CCGGCGCTCC CGCAGCCAGG AGCGGCGGGA GATGACGCGC GAGATCCTCT 300
 CCATTTTGGG CTGTCCCCAC CGCCCGCGCC CGCACCTCCA GGGCAAGCAC AACTCGGCAC 360
 CCATGTTTAT GCTGACCTG TACAAGCCCA TGGCGGTGGA GAGGGGCGGC GGGCCGGCGG 420
 20 GCCAGGCGTT CTCCTACCCC TACAAGSCCG TCTTCAGTAC CCAGGGCCCC CCTCTGGCCA 480
 GCTCGAAGA TAGCCATTTC CTCACCGAGC CCGACATGGT CATGAGCTTC GTCAACTCTG 540
 TGGAACTGA CAAGGAATTC TTCCAACCCG GCTACCAACA TCGAGAGTTC CGGTTTGATC 600
 TTTCACAGAT CCCAGAAAGG GAAGCTGTCA CGGCAGCCGA ATTCGGATC TACAAGGACT 660
 ACATCCGGGA AGCTTCCGAC AATGAGACGT TCCGGATCAG CGTTTATCAG GTGCTCCAGG 720
 AGCACTTGGG CAGGGAATCG GATCTCTTCC TGCTCGACAG CCGTACCCCT TGGGCTCCGG 780
 25 AGGAGGGCTG GCTGGTGTCT GACATCACAG CCACAGCAA CCAGTGGGTG GTCAATCCGC 840
 GGCACAACCT GGGCCTGAGC CTCTCGGTGG AGACGCTGGA TGGGACAGAG ATCAACCCCA 900
 AGTTGGGGGG CTTGATTTGG GGGCACGGGC CCCAGAACAA CGAGCCCTTC ATGGTGGCTT 960
 TCTTCAAGGC CACGAGGTTC CACTTCCGCA GCATCCGCTC CAGGGGAGC AACACGCGCA 1020
 GCCAGAACCG CTCGAGAGCG CCCAAGAAC AGGAAGCCCT GCGATGGGCC AACCTGGCAG 1080
 30 AGAACAGCAG CAGCGAGCAG AGGCAGGCTT GTAAGAAGCA CGAGCTGTAT GTCACTTCC 1140
 GAGACCTGGG CTGGCAGGAC TGGATCATCG CGCTGAAAG CTACGCGGCC TACTACTGTG 1200
 AGGGGGAGTG TGCCTTCCTT CTGAACCTCT ACATGAACGC CACCAACCA GCCATCGTGC 1260
 AGACGCTGGT CCACCTCATC AACCCGAAA CGGTGCCCAA GCGCTGCTGT GCGCCACGCG 1320
 AGCTCAATGC CATCTCCGTC CTCTACTTCC ATGACAGCTC CAACGTCATC CTGAAGAAAT 1380
 35 ACAGAAACAT GGTGGTCCGG GCGTGTGGCT GCGACTAGCT CCGCGAGAA TTCAGACCC 1440
 TTGGGGCCAA GTTTTCTGAG ATCTCTCATT GCTCGCCTTG GCCAGGAACC AGCAGACCAA 1500
 CTGCGTTTGG TGAGACCTTC CCTCCCTAT CCGCAACTTT AAAGGTGTGA GAGTATTAGG 1560
 AAACATGAGC AGCATATGGC TTTTGATCAG TTTTTCAGTG GCAGCATCCA ATGAACAAGA 1620
 TCTTACAAGC TGTGAGGCA AAACCTAGCA GGAAGAAAAA ACAACGCATA AAGAAAAATG 1680
 40 GCCGGGCGAG GTCATGTGCT GGGAAAGTCT AGCCATGCAC GGACTCGTCT CCAGAGGTAA 1740
 TTATGAGGCG CTACAGGCA GGGCAACCCG CCGTGGGAGG AAGGGGCGGT GGCAAGGGGT 1800
 GGGCACATTG GTGCTGTGCG GAAAGAAAAA TTGACCCGGA AGTTCTGTGA ATAAATGTCA 1860
 CAATAAACG AATGAATG

Seq ID NO: 659 Protein sequence
 Protein Accession #: NP_001710

1 11 21 31 41 51
 50 MHVRSRLAAA PHSFVALWAP LPLRLSALAD FSLDNEVHSS FIHRLRSQRE RREMQREILS 60
 ILGLPHRRPR HLQGGKHSAP MFMLDLYNAM AVEEGGGPGG QGFSPYKAV PSTQGPPLAS 120
 LQDSHFLTDA DMVMSFVNLV EHDKEFFHPR YHHRFRFDL SKIPEGEAVT AAEFRYKYDY 180
 IRERFDNETF RISVQVQLQE HLGRESDLFL LDSRTLWASE EGMIVFDITA TSNHWVNNPR 240
 55 HNLGLQLSVE TLGGQSINPK LAGLIGRHGP QNKQPFMVAP FKATEVHFRS IRSTGSKQRS 300
 QNRSKTPINQ EALRMANVAE NSSSDQRQAC KKHLYVSVFR DLGWQDWIIA PEGYAAYYCE 360
 GECAPFLNSY MNATNHAIVQ TLVHFINPET VPKPCCAPTQ LNAISVLYFD DSSNVILKKY 420
 RNMVVRACCG H

Seq ID NO: 660 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 211..1895

1 11 21 31 41 51
 65 GGATCTGAGG GGGCGCCAGT CACTTCTCTC ACGTTCTCGT GCTGGGCGGG AGGAGCGGAT 60
 GGGGCTTGGG AGGCAGCCTG CTCTCCAGTC CCTATCCACC CACAGGTTTT TTGGGTGGGA 120
 GAGGAATTAT CTGATAAAAT TCCGCGGTTA ATATTTTTAA AAACGGAGAG TTTTAAAAA 180
 TGATTTTTTT CCGTCGAAAA TGACCTTTTT ATGCTTCGAA GCAGTTTGTC AACCAGCATA 240
 70 GTGCTTTTTT TTTTCTCTTC TTTTCTACG ATAAATGAAA GCATTTCTTC AAGAAAAAGG 300
 CACAGGTTCC TTGAACAGCT GGATTCTGAT GGCACCATTA CTATAGAGGA GCAGATTGTC 360
 CTTGTGCTGA AAGCGAAAGT ACAATGTGAA CTCAACATCA CAGCTCACT CCAGAGGGGA 420
 GAAGGTAATT GTTTCCTGTA ATGGGATGGA CTCAATTGTT GGGCCAGAGG AACAGTGGG 480
 AAAAATACCG GTTGTCCATG CCTCCTTAT ATTATGACT TCAACCATAA AGGAGTTGCT 540
 75 TTCCGACACT GTAAACCCAA TGGAAACATG GATTTTATGC ACAGCTTAAA TAAACATAGG 600
 GCCAATTATC CAGACTGCCT TCGCTTCTG CAGCCAGATA TCAGCATAGG AAGCAAGAA 660
 TTCTTTGAAC GCGCTATGAT AATGTATACC GTTGGCTACT CCATCTCTTT TGGTTCCTTG 720
 GCTGTGGCTA TTCTCATCAT TGGTACTTTC AGACGATTGC ATTGCACTAG GAACTATATC 780
 CACATGCTAT TATTGTGTC TTTCATGCTG AGAGCTACAA GCATCTTTGT CAAGACAGA 840
 GTAGTCCATG CTCACATAGG AGTAAAGGAG CTGGAGTCCC TAATAATGCA GGATGACCCA 900
 80 CAAAATCTTA TTGAAGCAAC TTCTGTGGAC AAATCACAAT ATATCGGGT CAAGATTGCT 960
 GTTGTGATGT TTATTACTT CCTGGCTACA AATTATTATT GGATCCTGGT GGAAGGTCTC 1020
 TACCTGCATA ATCTCATCTT TGTGGCTTTC TTTTCGGACA CCAATACCT GTGGGGCTTC 1080
 ATCTTGATAG GCTGGGGGTT TCCAGCAGCA TTTGTGCGAG CATGGGCTGT GGCACGAGCA 1140
 ACTCTGGCTG ATCGCAGGTG CTGGGAACTT AGTGCTGGAG ACATCAAGTG GATTATCAA 1200
 85 GCACGATCT TAGCAGCTAT TGGCTGAAT TTTATTCTGT TTCTGATAC GGTAGAGTT 1260
 CTAGCTACCA AAATCTGGGA GACCAATGCA GTTGGGCTAG ACACAAGGA GCAATACAGG 1320
 AAACGGCCA AATCGACACT GGTCTGTGTC CTAGTCTTTG GAGTGCATTA CATGTTGTC 1380

5
10
15
20
25

GTATGCCTGC	CTCACTCCTT	CACTGGGCTC	GGGTGGGAGA	TCCGCATGCA	CTGTGAGCTC	1440
TTCTTCAACT	CCCTTCAGGG	TTTCTTTGTG	TCTATCATCT	ACTGCTACTG	CAATGGAGAG	1500
GTTTCAGGAG	AGGTGAAGAA	GATGTGGAGT	CGGTGGGAATC	TCTCCGTGGA	CTGGAAGAGG	1560
ACACCCCAT	GTGGCAGCG	CAGATGCGGC	TCAGTGCTCA	CCACCGTGAC	GCACAGCACC	1620
AGCAGCCAGT	CACAGGTGCG	GGCCAGCACA	CGCATGTGTC	TTATCTCTGG	CAAAGCTGCC	1680
AAGATCGCCA	GCAGACAGCC	TGACAGCCAC	ATCATTCTTAC	CTGGCTATGT	CTGGAGTAAC	1740
TCAGAGCAGG	ACTGCCCTGCC	ACACTCTTTT	CACGAGGAGA	CCAAGGAAGA	TAGTGGGAGG	1800
CAGGAGATG	ATATTCTAAT	GGAGAAGCCT	TCCAGGCCTA	TGGAATCTAA	CCCAGACACT	1860
GAAGGATGCC	AAGGAGAAAC	TGAGGATGTT	CTCTGA			

Seq ID NO: 661 Protein sequence
Protein Accession #: Eos sequence

15
20
25

1	11	21	31	41	51	
MLRSSLSISI	VLFLPSSPST	INESISSRKR	HRFLEQLDSD	GTITIEBQIV	LVLKAKVQCE	60
LNITAIQLQEG	EGNCFPEWDG	LICWPRGTVG	KISAVPCPPY	IYDFNHKGVA	FRHCNPNGTW	120
DPHSLNKTW	ANYSDCLRFL	QPDISIGKQE	PFERLYVMYT	VGYSISFGSL	AVAILIIGYF	180
RRLLHCTRNYI	RMHLFVSFML	RATSI FVKOR	VVAHIGVKEL	ESLIMQDDPQ	NSIEATSVDK	240
KSQYIGCKIA	VVMFIYFLAT	NYWILVEGL	YLHNLFVAF	FSDTKYLWGF	ILIGWGFPA	300
FVAHAWARA	TLADARCWEL	SAGDIKHIYQ	APILAAIGLN	FILPLNTVRV	LATKIWNETA	360
VGHDRKQYR	KLAKSTLVIV	LVFGVHYIVF	VCLPHSFTGL	GWEIRMHCEL	FPNSFQGFV	420
SIYCYCNGE	VQAEVKKMWS	RWNLSVDWKR	TPPGSRRCG	SVLITVTBST	SSQSQVAAST	480
RMVLISGKAA	KIASRQPDHS	ITLPGYVWSN	SEQDCLPHSF	HEETKEDSGR	QGDILMEKP	540
SRPMESNPDT	EGCGGETEDV	L				

Seq ID NO: 662 DNA sequence
Nucleic Acid Accession #: NM_005048
Coding sequence: 143..1795

30
35
40
45
50
55
60
65
70
75
80
85

1	11	21	31	41	51	
GGCCGGTGGC	CCGGGCCCGA	CCACCCAGC	TGCGCGTCGT	TACTGGCCAC	AAGTTTGCTC	60
TGGGCCAGCC	AAGTTGGCAA	CTTGGGAAGCT	TCTCCCGGGC	TCTGGAGGAG	GGTCCCTGCT	120
TCCTTCTACA	GCCTTCCCG	GCATGGCCGG	GCTGGGGGCG	TGCTTCCACG	TCTGGGGTTG	180
GCTAATGCTC	GGCAGCTGCC	TCCTGGCCAG	AGCCAGCCTG	GATTCTGATG	GCACCATTAC	240
TATAGAGGAG	CAGATTGTCC	TTGTGCTGAA	AGCGAAAGTA	CAATGTGAAC	TCAACATCAC	300
AGCTCAACTC	CAGGAGGAG	AAGGTAATTG	TTTCCCTGAA	TGGGATGGAC	TCAATTGTTG	360
GCCCGAGGGA	ACAGTGGGGA	AAATATCGGC	TGTTCCATGC	CCTCCTTATA	TTTATGACTT	420
CAACCATAAA	GGATTGTCTT	TCCGACACTG	TAAACCCCAAT	GGAAACATGG	ATTTTATGCA	480
CAGCTTAAT	AAAACATGGG	CCAATTATTC	AGACTGCCTT	CGCTTCTGCG	AGCCAGATAT	540
CAGCATAGTA	AAGCAAGAA	TCTTTGAAAC	CCTCTATGTA	ATGTATACCG	TTGGCTACTC	600
CATCTCTTTT	GGTTCCTTGG	CTGTGGCTAT	TCTCATCATT	GGTACTTCA	GACGATTGCA	660
TTGCACATAG	AACATATACC	ACATGCACTT	ATTTGTGTCT	TTATGTCTGA	GAGCTACAAG	720
CATCTTTGTC	AAAGACAGAG	TAGTCCATGC	TCACATAGGA	GTAAGGAGAG	TGGAGTCCCT	780
AATAATGCA	GATGACCCAC	AAATTTCCAT	TGAGGCAACT	TCTGTGAGCA	AATCACATAA	840
TATCGGGTGC	AAGATTGCTG	TTGTGATGTT	TATTTACTTC	CTGGCTACAA	ATTATTATTG	900
GATCCTGGTG	GAAGGTCTCT	ACCTGCATAA	TCTCATCTTT	GTGGCTTTCT	TTTCCGACAC	960
CAATACCTTG	TGGGGCTTCA	TCTTGATAGG	CTGGGGGTTT	CCAGCAGCAT	TTGTGTCAGC	1020
ATGGGCTGTG	GCACGACCAA	CTCTGGCTGA	TGCGAGGTGC	TGGGAACCTA	GTGCTGAGAG	1080
CATCAAGTGG	ATTTATCAAG	CACGATCTT	AGCAGCTATT	GGGCTGAATT	TTATCTGTGT	1140
TCGAATAAG	GTTAGAGTTC	TAGCTACCAA	AATCTGGGAG	ACCAATGCAG	TTGGGCATGA	1200
CACAAGGAAG	CAATACAGGA	AACTGGCCAA	ATGACACTG	GTCTGGTTC	TAGTCTTTGG	1260
AGTCATTAAC	ATCTGTTCG	TATGCCTGCC	TCATCTCTTC	ACTGGGCTCG	GGTGGGAGAT	1320
CGCATGACAC	TGTGAGCTCT	TCTTCAACTC	CTTTCAGGGT	TTCTTTGTGT	CTATCATCTA	1380
CTGCTACTGC	AATGGAGAGG	TTGAGGCAGA	GGTGAAGAAG	ATGTGGAGTC	GGTGAATCT	1440
CTCCGTGGAC	TGGAAAAGGA	CACGCCCATG	TGGCAGCCGC	AGATGCGGCT	CAGTGTCTAC	1500
CACCGTGAGC	CACAGCACCA	GCAGCCAGTC	ACAGGTGGCG	GCCAGCACAC	GCATGGTGCT	1560
TATCTCTGCG	AAAGCTGCCA	AGATGCCAG	CAGACAGCCT	GACAGCCACA	TCATCTTACC	1620
TGGCTATGTC	TGCTGTAATC	CAGAGCAGGA	CTGCCCTGCCA	CACCTCTTCC	ACGAGGAGAC	1680
CAAGGAAGAT	AGTGGGAGGC	AGGAGATGA	TATTCTAATG	GAGAGCCTT	CCAGGCTTAT	1740
GGAACTTAAC	CCAGACACTG	AAGGATGCCA	AGGAGAAACT	GAGGATGTTT	TCGTAATGGA	1800
CATTGTGTGC	TGACTTTTAT	GGGCTGGTCC	AATGGCTGGT	TGTGTGAGAG	GGCTTGGCTG	1860
ATACTCCTAT	CGTGTAGTTC	AAAGGCTGAA	AATTCAGTTA	AGGTGTACT	TAATAATAGT	1920
TTTTAGGCTC	CATGAATTGG	CTCCTGTAAA	TACTAACGAC	ATGAAAATGC	AAGTGTCAAT	1980
GGAGTAGTTT	ATTACCTTCT	ATTGGCATCA	AGTTTCTCCT	TAAATTAATG	TATGTTATTT	2040
GCTCTGTGAT	TGTTTCAATTT	TTTCTGTCTAC	TTTTGGGTAG	AAAAAAGATT	CAATTGCTTG	2100
GCTGTAGCTT	TCTCTCATAT	ATATCACCTT	AAATATAATG	AAGATCTTTT	AGTGTGTATC	2160
ATTTTCTTTT	TAGAACTAG	TATTTCTTTA	TTTCTTACTT	TAATGTACTT	CTATCAGTGC	2220
ATTTATTTTG	CCTGTGCATA	GGAGCAATTA	GGATCTAAAA	AAATATATGG	GAAGATAAAA	2280
GATCTAAGAA	CAAGTACTTG	CTGGAAAATT	AGTTGGCTGG	ACATTGTATA	AATAATGCAT	2340
TTATAACAAT	TACATGTGTT	TTTGGGAACA	AGGAAAAATT	CTCAAAAAG	AATATTTTAC	2400
ACATCCCTTC	TTTTGAATGG	CCTCTTTGTG	ACCAGCCAGA	CCTCAGGTCT	TCATCTTTTC	2460
TTCTTTGTAA	ACCATGTCAT	GTGGAAAAGT	TTCTCAGTTT	AGTGAGCTTG	TGCTGCAAAA	2520
TTGATTTTGT	TTTGAATGTA	TTTTGATAGC	AAATCATGCT	GCATCTATAT	CTTTTCTTGT	2580
TTTGAGCTGT	TACTACATTG	TACATGGCAT	GTGGGATCAA	TAAAAAATTT	GTTTAAAAAA	2640

Seq ID NO: 663 Protein sequence
Protein Accession #: NP_005039

80
85

1	11	21	31	41	51	
MAGLGASLHV	NGWMLGSCIL	LARAQLSDG	TITIEBQIVL	VLKAKVQCEL	NITAIQLEGE	60
GNCFPEWDGL	ICWPRGTVGK	ISAVPCPPYI	YDFNHKGVA	RHCNPNGTWD	FMHSLNKTWA	120
NYSDCLRFLQ	PDISIGKQEF	FERLYVMYT	VGYSISFGSLA	VAILIIGYPR	RLHCTRNYIH	180
MHLFVSFMLR	ATSI FVKORV	VVAHIGVKEL	ESLIMQDDPQ	NSIEATSVDK	SQYIGCKIAV	240

VMPYFLATN YYNIVVEGLY LHNLPVAFV SDTKYLWGF I LIGWGFPAF VAAWAVARAT 300
 LADARCWELS AGDIKWIYQA PILAAIGLNF ILPLNTVRVL ATKIWETNAV GHDTRKQYRK 360
 LAKSTLVVLV VFGVHYIVFV CLPHSFTGLG WEIRMHCELF FNSPQGFVFS IYCYCNGEV 420
 QAEVKQWMSR WNLSDVWKRT PPGSRRRCGS VLTITVTHSTS SQSQVAASR MVLISGKAOK 480
 IASRQPDHSI TLPQVWSNS EQDCLPHSFB EETKEDSGRQ GDDILMEKPS RPMESNPDTE 540
 GCGGETEDVL

Seq ID NO: 664 DNA sequence
 Nucleic Acid Accession #: NM_012152
 Coding sequence: 43..1104

1 11 21 31 41 51
 CTCTCTTAAA TTCTCTTCTA GGATGTTTAC TTCTCTTCCA CAATGAATGA GTGTCACTAT 60
 GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTGCA TGACTGGACA 120
 GGAAACAAGC TTGTGATTGT TTGTGTGTGT GGGACGTTTT TCTGCGCTGT TATTTTTTTT 180
 TCTAATTTCT TGTGTATCGC GGCAGTGATC AAAACAGAAA AATTTCAATT CCCCTTCTAC 240
 TACCTGTGGG CTAATTAGC TGCTGCCGAT TTCTGCGCTG GAATTGCTTA TGTATTCTGT 300
 ATGTTTAAAC CAGGCCCACT TTCAAAAAC TTGACTGTCA ACGCTGCTT TCTCGTCTAG 360
 GGGCTTCTGG ACAGTGTACT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCGGTGGAG 420
 AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA 480
 CTGCTCATTT TGCTGTCTGT GGCATCGCC ATTTTATG GGGCGGTCCC CACACTGGGC 540
 TGGAAATGCC TCTGCAACAT CTCTGCCCTG TCTTCCCTGG CCCCATTTA CAGCAGGAGT 600
 TACCTTGT TTCTGACAGT GTCCAACCTC ATGGCCCTTC TCATCATGCT TGTGGTGTAC 660
 CTGCGGATCT ACCTGTACTG CAAGAGGAAA ACCAAGCTCT TGTCTCGCA TACAAGTGGG 720
 TCCATCAGCC CGCGGAGGAC ACCCATGAAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG 780
 GGGTTTGTGG TATGCTGGAC CCGGGGCTGT GTGGTCTGCT TCCTCGACGG CTGAACTGC 840
 AGGCAGTGTG GGTGTGAGCA TGTGAAAAGG TGGTTCTGCT TGCTGGCGCT GCTCAACTCC 900
 GTGGTGAACC CCATCATCTA CTCTACAAG GAGCAGGACA TGTATGGCAC CATGAAGAAG 960
 ATGATCTGCT GCTTCTCTCA GAGAAACCCA GAGAGCGCTC CCTCTGCGAT CCCCTCCACA 1020
 GTCTCAGCA GAGTGTACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080
 GTCTGCAATA AAGCACTTC CTAACCTCTG GATGCTCTCT GGGCCACCCA GGTGATGACT 1140
 GTCTTAGG

Seq ID NO: 665 Protein sequence
 Protein Accession #: NP_036284

1 11 21 31 41 51
 MNECHYDKHM DFFYNRSNTD TVDDWTGTLK VIVLCVGTFF CLPIFFSNLS VIAAVIKNRK 60
 FHFFFFYLLA NLAAADFFAG IAYVFLMFT GPVSKTLTVN RWFLRQGLLD SSLTASLTNL 120
 LVIAVERHMS IMMRVHNSL TKKRVTLILL LVWAIAPMG AVPTLGNCL CNISACSSLA 180
 PIYSRSLVFP WTVSNLWAPL IMVVVYLRIV VYVKRKTNL SPHTSGSISR RRTPMKLMKT 240
 VMTVLGAPVW CWTPLGVLL LDGLNCRQCG VQHKRWFLLL LALLNSVNVN IYYSKDEDM 300
 YGTMKQMIC PSQENFERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS

Seq ID NO: 666 DNA sequence
 Nucleic Acid Accession #: NM_002821
 Coding sequence: 150..3362

1 11 21 31 41 51
 AACTCCCGCC TCGGGACGCC TCGGGGTGCG GCTCCGGCTG CGGCTGCTGC TGGCGGCCCC 60
 GCGCTCGGTG GCGTCCGCTC CCGTGTGCGG CCGGGAGGCA GTCTGCGGCC CGCGGTGCGC 120
 CCTCAGCTCC TTTTCTGAG CCGCGCGGCA TGGGAGCTGC GCGGGGATCC CCGGCCAGAC 180
 CCGCGCGGTT GCGTCTGCTC AGCGTCTGCT TGCTGCGCTG GCTGGGCGGT ACCCAGACAG 240
 CCATTGTCTT CATCAAGCAG CCGTCTCTCC AGGATGCACT GCAGGGGCGC CGGGCGCTGC 300
 TTGCTGTGTA GGTGTAGGCT CCGGGCCCGG TACATGTGTA CTGGCTGCTC GATGGGGCCC 360
 CTGTCCAGGA CAGCGAGCGG CGTTTCCGCC AGGGCAGCAG CCTGAGCTTT GCAGCTGTGG 420
 ACCGGCTGCA GGAATCTGCG ACCTTCCAGT GTGTGGCTCG GGATGATGTC ACTGGAGAAG 480
 AAGCCCGCAG TGCACAGGCC TCCTTCAACA TCAAATGGAT TGAGGCAGGT CCTGTGGTCC 540
 TGAAGCATCC AGCCTCGGAA GCTGAGATCC AGCCACAGAC CCAGGTCAAC CTTGTTGCCC 600
 ACATTGATGG GCACCTCGG CCCACCTACC AATGTTCCG AGATGGGACC CCCCTTTCTG 660
 ATGGTCAGAG CAACCAACA GTCAAGCAGCA AGGAGCGGAA CCTGACGCTC CGGCCAGCTG 720
 GTCTGAGCA TAGTGGGCTG TATTCTGCT GCGCCACAG TGCTTTTGGC CAGGCTTGCA 780
 GCAGCCAGAA CTTCACTTGT AGCATTGCTG ATGAAAGCTT TGCCAGGGTG GTGCTGGCAC 840
 CCCAGGAGCT GGTAGTAGCG AGGTATGAGG AGGCCATGTT CCATTGCCAG TTCTCAGCCC 900
 AGCCACCCCC GAGCTGCGAG TGGCTCTTTG AGGATGAGAC TCCCATCACT AACCGCAGTC 960
 GCGCCGCCCA CCTCCGAGA GCCACAGTGT TTGCCAACGG GTCTCTGCTG CTGACCCAGG 1020
 TCCGGCCAGC CAATGCAAGG ATCTACCGCT GCATTGGCCA GGGGCGAGAG GGGCCACCCA 1080
 TCATCTCGGA AGCCACATT CACTAGCAG AGATTGAAGA CATGCCGCTA TTTGAGCCAC 1140
 GGGTGTTTAC AGCTGCGAGC GAGGAGCGTG TGACCTGCCT TCCCCCAAG GGTCTGCCAG 1200
 AGCCAGAGCT CTGGTGGGAG CAGCGGGGAG TCCGGCTGCC CACCCATGGC AGGGTCTACC 1260
 AGAAGGGCCA CGAGCTGGTG TTGGCCAATA TTGCTGAAAG TGATGCTGGT GTCTACACCT 1320
 GCGACGCGGC CAAGCTGGCT GGTCAAGCGA GACAGGATGT CAACATCACT GTGGCCACTG 1380
 TGCCCTCTCT GCTGAAGAAG CCCCAAGACA GCCAGCTGGA GGAGGGCAAA CCGGCTACT 1440
 TGGATTGCTT GACCCAGGCC ACACCAAAAC CTACAGTTGT CTGGTACAGA AACCCAGATG 1500
 TCATCTCAGA GGAATCAAGG TTGAGGCTCT TCAAGATGG GACCTTGGCG ATCAACAGCG 1560
 TGGAGTGTGA TGTATGGGCA TGGTACCGTT GTATGAGCAG CACCCAGGCC GGCAGCATCG 1620
 AGGCGCAGC CCGTGTCCAA GTGCTGGAAG AGCTCAAGTT CACACCAACA CCCCAGCCAC 1680
 AGCAGTGCAAT GGAGTTTGAC AAGGAGGCCA CGGTGCCCTG TTCAGCCACA GGGCGAGAGA 1740
 AGCCCACTAT TAAGTGGGAA CGGGCAGATG GGAGCAGCCT CCCAGAGTGG GTGACAGACA 1800
 ACGCTGGGAC CTTGCAATTT GCCCGGTGA CTGAGATGA CCGTGGCAAC TACACTTGCA 1860
 TTGCTCCAA CGGGCCGAGG GGCCAGATTG GTGCCCATGT CCAGCTCACT GTGGCAGTTT 1920
 TTATCACTTT CAAAGTGGAA CCAGAGGCTA CGACTGTGTA CCAGGGCCAC ACAGCCCTAC 1980
 TGCAGTGGCA GGGCCAGGGG GACCCCAAGC CGCTGATTCA GTGGAAGGCG AAGGACCGCA 2040
 TCCTGGACCC CACCAAGCTG GGACCCAGGA TGACATCTTT CCAGATGGC TCCCTGGTGA 2100

	TCCATGACGT	GGCCCTGAG	GACTCAGGCC	GCTACACCTG	CATTGCAGGC	AACAGCTGCA	2160
	ACATCAAGCA	CACGGAGGCC	CCCTCTATG	TGCTGGACAA	GCTGTGCGG	GAGGAGTCGG	2220
	AGGGCCCTGG	CAGCCCTCCC	CCCTACAAGA	TGATCCAGAC	CATTGGGTTG	TGGTGGGTG	2280
5	CCGCTGTGGC	CTACATCAT	GCCGTGCTGG	GCTTCATGTT	CTACTGCAAG	AAGCGCTGCA	2340
	AAGCCAGCG	GCTCGAGAAG	CAGCCCGAGG	GCGAGGAGCC	AGAGATGGAA	TGCTCAACG	2400
	GAGGGCTTT	GCAGAACGGG	CAGCCCTCAG	CAGAGATCCA	AGAAGAAGTG	GCCTTGACCA	2460
	GCTTGGGCTC	CGGCCCCGCG	GCCACCAACA	AACGCCACAG	CACAAGTGAT	AAGATGCACT	2520
	TCCCACGGTG	TAGCCTGCG	CCCATCACCA	CGCTGGGGAA	GAGTGAGTTT	GGGGAGGTGT	2580
10	TCCTGCAAA	GGCTCAGGGC	TTGGAGGAGG	GAGTGGCAGA	GACCTTGGTA	CTTGTGAAGA	2640
	GCCTGCGAG	GAGGATGAG	CAGCAGCAGC	TGGACTTCCG	GAGGGAGTTG	GAGATGTTTG	2700
	GGAGCTGAA	CCACGCCAAC	GTGGTGGCGC	TCCTGGGGCT	GTGCCGGGAG	GCTGAGCCCC	2760
	ACTACATGGT	GCTGGAATAT	GTGGATCTGG	GAGACCTCAA	GCAGTTCTCT	AGGATTTCCA	2820
	AGAGCAAGGA	TGAAAAATTG	AAGTCACAGC	CCCTCAGCAC	CAAGCAGAAG	GTGGCCCTAT	2880
	GCACCCAGAT	AGCCCTGGGC	ATGGAGCACC	TGTCCAACAA	CCGCTTTGTG	CATAAGGACT	2940
15	TGGCTGCGCG	TAATGCTGCT	GTGATGCCCC	AGAGACAAGT	GAAGGTGTCT	GCCTTGGGCC	3000
	TCAGCAAGGA	TGTGTACAAC	AGTGAGTACT	ACCCTTCCCG	CCAGGCGCTGG	GTGCGCTGCG	3060
	GCTGGATGTC	CCCCGAGGCC	ATCCTGGAGG	GTGACTTCTC	TACCAAGTCT	GATGCTCTGG	3120
	CCTTGGGTGT	GCTGATGTGG	GAAAGTGTTA	CACATGGAGA	GATGCCCAT	GSTGGGCGAG	3180
20	CAGATGATA	AGTATGGCA	GATTGTCAGG	CTGGGAAGGC	TAGACTTCTC	CAGCCCGAGG	3240
	GCTGCCCTTC	CAAACTCTAT	CGGCTGATGC	AGCGCTGCTG	GGCCCTCAGC	CCCAAGGAAC	3300
	GGCCCTCCTT	CAGTGAGATT	GCCAGCGCCC	TGGGAGACAG	CACCGTGGAC	AGCAAGCCGT	3360
	GAGGAGGGAG	CCGCTCAGG	ATGGCCTGGG	CAGGGGAGGA	CATCTCTAGA	GGGAGAGCTCA	3420
	CAGCATGATG	GGCAAGATCC	CTGCTCTCCT	GGGCCCTGAG	GTGCCCTAGT	GCAACAGGCA	3480
25	TTGCTGAGGT	CTGAGCAGGG	CCTGGCCTTT	CCTCCTCTTC	CTCACCCCTCA	TCCCTTGGGA	3540
	GGCTGACTTG	GACCCAAACT	GGGCGACTAG	GGCTTTGAGC	TGGGCAGTTT	CCCTGCGCAC	3600
	CTCTTCTCTT	ATCAGGGACA	GTGTTGGTGC	CACAGGTAAC	CCCAATTTCT	GGCCTTCAAC	3660
	TTCTCCCTCT	GACCGGCTCC	AACTCTGCCA	CTCATCTGCC	AACCTTGGCT	GGGAGGGGCT	3720
	AGGCTTGGGA	TGAGCTGGGT	TTGTGGGGAG	TTCTTAAATA	TTCTCAAGTT	CTGGGCACAC	3780
30	AGGGTAAAGT	AGTCTCTTGC	CCACTGGTCC	ACTTGGGGGT	CTAGACCAGG	ATTATAGAGG	3840
	ACACAGCAAG	TGAGTCTCTC	CAACTCTGGG	CTTGTGCACA	CTGACCAGGA	CCCACTCTCT	3900
	CCCCACCTTT	CTCTCTCTTC	CTCATCTCTA	GTGCTTGGCA	GATGAAGGAG	TTTTCAGGAG	3960
	CTTTTGACAC	TATATAAAC	GCCCTTTTGT	TATGCACCAC	GGCGCGCTTT	TATATGTAAT	4020
	TGACAGCGTG	GGTGGGTGGG	CATGGGAGGT	AGGGGTGGGC	CCTGGAGATG	AGGAGGGTGG	4080
35	GCCATCCTTA	CCCCACACT	TTATTGTTGT	CGTTTTTTGT	TTGTTTTGTT	TTTTTGTTTT	4140
	TGTTTTTGT	TTTACACTCG	CTGCTCTCAA	TAAATAAGCC	TTTTTTA		

Seq ID NO: 667 Protein sequence
Protein Accession #: NP_002812

	1	11	21	31	41	51	
	MGAARGSPAR	PRRLPLLSVL	LLPLLGGTQT	AIIVFIKQPS	QDALQRRAL	LRCVEAPGP	60
	VHVYWLDDGA	PRVQDTERFA	QGSSLSFAAV	DRLQDSGTFO	CVARDVDTGE	EARSANASFN	120
	IKWIEAGPVG	LKHPSAEAEI	QPQTQVTLRC	HIDGHPRTY	QWFRDGTPLS	DGQSNHTVSS	180
45	KERNLTLRPA	GPEHSLGYSC	CAHSAPQOAC	SSQNFTLSIA	DESPARVULA	PQDVVVARYE	240
	EAMFHQFSA	QPPPSLQWLF	EDETPIITNRS	RPHHLRRATV	FANGSLLLTQ	VRPRNAGIYR	300
	CIGQGGRGPP	IILEATLHLA	EIEDMPLEFEP	RVFTAGSEER	VTCLFPKGLP	EPSVWWEHAG	360
	VRLEPTGRVY	QKGEHLVLAN	IAESDAGVYT	CHAANLAGOR	RQDVNITVAT	VPSWLKQPQD	420
	SQLBEGKPGY	LDCITQATPK	PTVVWYRNQM	LISEDSRFEV	FINGTLRINS	VEVYDGTWYR	480
50	CMSSTPAGSI	EAQARVQVLE	KLKFTPPPPQ	QOCMEFDKEA	TVPCSATGRE	KPTIKWERAD	540
	GSSLPBWVTD	NAGTLHFARV	TRDDAGNYTC	IASNGPQQGI	RAHVQLTVAV	FITPKVEPER	600
	TTVYQGHIAL	LQCEAQGDPK	PLIQWKGKDR	ILDPTKLGRP	MHIFQNGSLV	IHDVAPEDSG	660
	RYTCLAGNRC	NIKTEAPLY	VVDKPVPEES	EGPGSPPPYK	MIQTIGLSVG	AAVAYIIAIV	720
	GLMFYCKKRC	KAKRLQKQPE	GEEPEMECLN	GGPLQNGQPS	AEIQEEVALT	SLGSGPAATN	780
55	KRHSTDKMHH	PRGSSSLQPT	TLGKSEFGEV	FLAKAQGLEE	GVAETLVLVK	SLQTKDEQQQ	840
	LDPRRELEMF	GKLNHANVVR	LLGLCREAEP	HYMVLEYVDL	GDLLQFLRIS	KSKDEKLKSK	900
	PLSTKQKVAL	CTQVALGMEH	LSNNRFVHKD	LAARNCLVSA	QRQVKVSALG	LSKDVYNSEY	960
	YHFRQAKVPL	RWMSPEAILL	GDFSTKSDVW	AFGVLMNEVF	THGEMPHGGQ	ADDEVLLADLQ	1020
60	AGKARLPQPE	GCPSKLYRLM	QRCHALSPKD	RPSFSEIASA	LGDSTVDSKP		

Seq ID NO: 668 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1389

	1	11	21	31	41	51	
	ATGGGCTACC	AGAGGCAGGA	GCCTGTATC	CGCCCGCAGA	GAGATTTAGA	TGACAGAGAA	60
	ACCCTTGTGT	CTGAACATGA	GTATAAAGAG	AAAACCTGTC	AGTCTGCTGC	TCTTTTAAAT	120
	GTGTCAACT	CGATTATAGG	ATCTGGTATA	ATAGGATTGC	CTTATTCAAT	GAGCAAGCT	180
70	GGGTTTCTCT	TGGGAATATT	GCTTTTATTC	TGGGTTTCAT	ATGTTACGGA	CTTTTCCCTT	240
	GTTTTATTGA	TAAAGGAGG	GGCCCTCTCT	GGAACAGATA	CCTACCAGTC	TTTGGTCAAT	300
	AAAACTTTCG	GCTTTCCAGG	GTATCTGCTC	CTCTCTGTTT	TTGATTTTCT	GTATCCTTTT	360
	ATAGCAATGA	TAGTTACAA	TATAATAGCT	GGAGATACCT	TGAGCAAAGT	TTTTCAAAGA	420
	ATCCAGGAG	TGATCCTGA	AAAAGTGTGT	ATTGGTGGCC	ACTTCATTAT	TGCACTTTCC	480
75	ACAGTTACCT	TTATCTGCCC	TTTATCCTTG	TACCGAAATA	TAGCAAAGCT	TGGAAAGGTC	540
	TCCCTCATCT	CTACAGGTTT	AACAACTCTG	ATTCTTGGAA	TGTAAATGGC	AAGGGCAATT	600
	TCACCTGGTC	CACATATACC	AAAAACAGAA	GACGCTTGGG	TATTTGCAAA	GCCCAATGCC	660
	ATTCAAGCGG	TGGGGGTTAT	GTCTTTTGCA	TTTATTTGCC	ACCATAACTC	CTTCTTAGTT	720
	TACAGTTCTC	TAGAAGAAGC	CACAGTAGCT	AAGTGGTCCC	GCCTTATCCA	TATGTCCATC	780
80	GTGATTCTGT	TATTATCTGT	TATATTCTTT	GCTACATGTG	GATACCTGAC	ATTTACTGGC	840
	TTCAACCAAG	GGGACTTATT	TGAAAATTAC	TGCAGAAATG	ATGACCTGGT	AACATTTGGA	900
	AGATTTTGTG	ATGGGTGTC	TGTCAATTTG	ACATACCTTA	TGGAATGCTT	TGTGACAAGA	960
	GAGGTAATTG	CCAATGTGTT	TTTTGGTGGG	AATCTTTTAT	CGSTTTTCCA	CATTGTTGTA	1020
	ACAGTGATGG	TCACTACTGT	AGCCACGCTT	GTGTCATTGC	TGATTGATGG	CCTCGGGATA	1080
85	GTTCTAGAAC	TCAATGGTGT	GCTCTGTGCA	ACTCCCTCCA	TTTTTATCAT	TCCATCAGCC	1140
	TGTTATCTGA	AACGTCTCTA	AGAACCAAGG	ACACACTCCG	ATAAGATTAT	GTCTTGTGTC	1200
	ATGCTTCCCA	TTGGTGCTGT	GGTGATGGTT	TTTGATTCG	TCATGGCTAT	TACAAATACT	1260

CAAGACTGCA CCCATGGGCA GGAAATGTTT TACTGCTTTC CTGACAATTT CTCTCTCACA 1320
AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACCTT CTACTTTAAA TATTAGTATC 1380
TTTCAATGA

5

Seq ID NO: 669 Protein sequence
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
10	MGYQRQEPVI	PPORDLDDRE	TLVSEHEYKE	KTCQSAALFN	VVNSIIGSGI	IGLPYSMKQA	60
	GFPLGILLLF	WVSIVTDFSL	VLLIKGGALS	GTDTYQSLVN	KTFGFGYLL	LSVLQFLYFP	120
	IAMISYNIIA	GDTLKVPQR	IPGVDPENVF	IGRHFIIGLS	TVTPTLPLSL	YRNIKLGKV	180
	SLISTGLTTL	ILGIVMARAI	SLGPHIPKTE	DAMVFAKPNA	IQAVGVMSFA	FICHENSFLV	240
15	YSSLEPTVA	KWSRLIHMSI	VISVFCIFP	ATCGYLTFTG	FTQGDLFENY	CRNDDLVTFG	300
	RFCYGVTVIL	TYPMCEFVTR	EVIANVFPGG	NLSSVPHIVV	TVMVITVATL	VSLLDICLGI	360
	VLELNGVICA	TPLIFIIPSA	CYLKLSSEPR	THSDKIMSCV	MLPIGAVVMV	PGFVMAITNT	420
	QDCTHQEMF	YCFPDNFSLT	NTSESHVQQT	TQLSTLWISI	PQ		

20

Seq ID NO: 670 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1284

	1	11	21	31	41	51	
25	ATGGGCTACC	AGAGGCAGGA	GCCGTGCATC	CCGCGCAGCA	GAGGATTGCC	TTATTCAATG	60
	AAGCAAGCTG	GGTTTCCCTT	GGGAATATTG	CTTTTATTCT	GGGTTTCATA	TGTTACAGAC	120
	TTTTCCCTTG	TTTTATTGAT	AAAAGGAGGG	GCCCTCTCTG	GAACAGATAC	CTACCACTCT	180
	TGGTCAATA	AAACTTTCGG	CTTTCAGGGG	TATCTGCTCC	TCTCTGTCTT	TCAGTTTTTG	240
	TATCCTTTTA	TAGCAATGAT	AAGTTACAAT	ATAATAGCTG	GAGATACTTT	GAGCAAAAGT	300
30	TTTCAAAGAA	TCCAGAGAGT	TGATCCTGAA	AACGTGTTTA	TGGTGCAGCA	CTTCATTATT	360
	GGACTTTCCA	CAGTTACCTT	TACTCTGCCT	TTATCCTTGT	ACCGAAATAT	AGCAAAGCTT	420
	GGAAAGTCT	CCCTCATCTC	TACAGGTTTA	ACAACTCTGA	TTCTTGAAT	TGTAATGGCA	480
	AGGCAATTT	CACGTGGTCC	ACACATACCA	AAAACAGAA	ACGCTTGGGT	ATTTGCAAA	540
	CCCAATGCCA	TCCAAGCGGT	CGGGTTATG	TCTTTTGCAT	TTATTGCGCA	CCATAACTCC	600
35	TTCTTAGTTT	ACAGTTCTCT	AGAAGAAGCC	ACAGTAGCTA	AGTGGTCCCG	CCTTATCCAT	660
	ATGTCCATCG	TGATTTCTGT	ATTATCTGT	ATATTCTTTG	CTACATGTGG	ATACCTGACA	720
	TTTACTGGCT	TCACCAAGGG	GGACTTATTT	GAAAATTACT	GCAGAAATGA	TGACCTGGTA	780
	ACATTGGGAA	GATTTTGTGA	TGGTGTCACT	GTCATTTTGA	CATACCCCTAT	GGAAATGCTT	840
	GTGACAAGAG	AGGTAATTGC	CAATGTGTTT	TTTGGTGGGA	ATCTTTCATC	GGTTTTCAC	900
40	ATTGTTGTAA	CAGTGATGGT	CATCACTGTA	GCCACGCTTG	TGTCATTGCT	GATTGATTGC	960
	CTCGGATAG	TTCTAGAACT	CAATGGTGTG	CTCTGTGCAA	CTCCCTCAT	TTTTATCAAT	1020
	CCATCAGCTG	GTTATCTGAA	ACTGTCTGAA	GAACCAAGGA	CACACTCCGA	TAAGATTATG	1080
	TCTGTGTCA	TGCTCCCAT	TGGTGTCTGT	GTGATGTTT	TTGGATTGCT	GATGGCTATT	1140
45	ACAAATCTC	AAGACTGCAC	CCATGGGCAG	GAAATGTTCT	ACTGCTTCC	TGACAATTC	1200
	TCTCTACAA	ATACCTCAGA	GTCTCATGTT	CAGCAGACAA	CACAACTTTC	TACTTTAAAT	1260
	ATTAGTATCT	TCAACTCGA	GTAA				

50

Seq ID NO: 671 Protein sequence
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
	MGYQRQEPVI	PPQRLPYSYM	KQAGFPLGIL	LLFWVSIVTD	FSLVLLIKGG	ALSGTDTYQS	60
	LVNKTGFPFG	YLLSLVLQPL	YPIAMISYN	IIAGDTLSKV	QRIPIGVDP	NVFIHRHPII	120
55	GLSTVTFTLP	LSLYRNIAKL	GKVSLLSTGL	TLILGIVMA	RAISLGPHIP	KTEDAWVFAR	180
	FNAIQAVGVM	SFAPICHENS	FLVYSSLEEP	TVAKWSRLIH	MSIVISVFIC	IFFATCGYLT	240
	FTGFTQGLDF	ENYCRNDDL	TVGRFCYGV	VILTYPMECF	VIREVIANVF	PGNLSVVFH	300
	IVTVVMTIV	ATVLSLLIDC	LGIVLELNGV	LCATPLIFII	PSACYLKLSE	EPRTSHDKIM	360
60	SCVMLPIGAV	VWVFGFVMAI	TNTQDCTHQ	EMFYCFPDNF	SLTNTSESHV	QQTQLSTLN	420
	ISIFQLE						

65

Seq ID NO: 672 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1203

	1	11	21	31	41	51	
	ATGGGCTACC	AGAGGCAGGA	GCCGTGCATC	CCGCGCAGT	TTCCCTTGT	TTATTGATA	60
	AAAGGAGGGG	CCCTCTCTGG	AACAGATACC	TACCAGICTT	TGGTCAATAA	AACITTCGGC	120
70	TTTCCAGGGT	ATCTGCTCCT	CTCTGTCTCT	CAGTTTTTGT	ATCCTTTTAT	AGCAATGATA	180
	AGTTACAATA	TAATAGCTGG	AGATACCTTG	AGCAAGATTT	TTCAAAGAAT	CCCAGGAGTT	240
	GATCCTGAAA	ACGTGTTTAT	TGGTGCACAC	TTCAATTATG	GACITTTCCAC	AGTTACCTTT	300
	ACTCTGCCCT	TATCCTTGTA	CCGAAATATA	GCAAAGCTTG	GAAAGGTCTC	CCTCATCTCT	360
	ACAGGTTTAA	CAACTCTGAT	TCTTGGAAAT	GTAATGGCAA	GGGCAATTTT	ACTGGGTCCA	420
75	CACATACCAA	AAACAGAAGA	CGCTTGGGTA	TTTGCAAAAG	CCAATGCCAT	TCAAGCGGTC	480
	GGGTTTATGT	CTTTTGCAAT	TATTTGCCAC	CATAACTCCT	TCTTAGTTTA	CAGTTCTCTA	540
	GAAGAACCBA	CAGTAGCTAA	GTTGTCCTCG	CTTATCCATA	TGTCCATCGT	GATTTCTGTA	600
	TTTATCTGTA	TATTTCTTGC	TACATGTGGA	TACTTGACAT	TTACTGGCTT	CACCAAGGG	660
	GACTTATTGT	AAAATTACTG	CAGAAATGAT	GACCTGGTAA	CATTTGGAAG	ATTTTGTAT	720
80	GGTGTACCTG	TCATTTTGAC	ATACCCATG	GAATGCTTTG	TGACAAGAGA	GGTAATTGCC	780
	AATGTGTTTT	TTGGTGGGAA	TCITTCATCG	GTTTTCACCA	TGTTTGAAC	AGTGATGGTC	840
	ATCAGTGTAG	CCACGCTTGT	GTCAATGCTG	ATTGATTGCC	TCGGGATAGT	TCTAGAAGTC	900
	AATGGTGTGC	TCTGTGCAAC	TCCCTCATTT	TTTATCATTC	CATCAGCCTG	TTATCTGAAA	960
	CTGTCTGAAG	AACCAAGGAG	ACACTCCGAT	AAGATTATGT	CTTGTGTCAT	GCCTCCCAT	1020
85	GGTGTGCTGG	TGATGGTTTT	TGGATTCTCT	ATGGCTATTA	CAAACTACTCA	AGACTGCACC	1080
	CATGGCAGG	AAATGTTCTA	CTGCTTTCCT	GACAAATTTCT	CTCTCACAAA	TACCTCAGAG	1140
	TCTCATGTTT	AGCAGACAAC	ACAACCTTCT	ACTTTAAATA	TTAGTATCTT	TCAACTCGAG	1200

TAA

Seq ID NO: 673 Protein sequence
Protein Accession #: Eos sequence

5

10

15

1	11	21	31	41	51	
MGYQRQEPVI	PPQFSLVLLI	KGGALSGTDT	YQSLVNKTGF	FPGYLLLSVL	QFLYPFIAMI	60
SYNIIAGDTL	SKVFPQIPGV	DPENVFIGRH	FIIGLSTVTF	TLPLSLYRNI	AKLGKVSLLS	120
TGLTTLILGI	VMARAILSGP	HIPKTEDAWV	FAKPNAIQAV	GVMSFAPICH	HNSFLVYSSL	180
EEPTVAKWSR	LIHMSIVISV	PICIFFATCG	YLTFGTGTQ	DLFENYCRND	DLVTFGRFCY	240
GTVILTITPM	ECFVTREIVIA	NVFFGGNLSS	VFHIVVTVMV	ITVATLVSL	IDCLGIVLEL	300
NGVLCATPLI	FIIPSACYLK	LSEEPRTSD	KIMSCVMLPI	GAVVMVFGPV	MAITNTQDCT	360
HGQEMFYCFP	DNFSLTNTSE	SHVQQTQLS	TLNISIFQLE			

Seq ID NO: 674 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1140

20

25

30

35

40

1	11	21	31	41	51	
ATGGGCTACC	AGAGGCAGGA	GCCGTGCATC	CCGCCGCAGG	TCAATAAAAC	TTTCGGCTTT	60
CCAGGGTATC	TGCTCTCTCT	TGTTCTTCAG	TTTTTGTATC	CTTTTATAGC	AATGATAAGT	120
TACAATATAA	TAGCTGGAGA	TACTTTGAGC	AAAGTTTTTC	AAAGAATCCC	AGGAGTTGAT	180
CCTGAAAGCG	TGTTTATGGG	TCCGCACCTC	ATTATTGGAC	TTTCCACAGT	TACCTTTACT	240
CTGCCCTTAT	CCTTGTACCG	AAATATAGCA	AAGCTTGGAA	AGGCTCTCCCT	CATCTCTACA	300
GGTTTAACAA	CTCTGATTC	TGGAAATGTA	ATGGCAAGGG	CAATTTCACT	GGGTCCACAC	360
ATACCAAAAA	CAGAAGACGC	TTGGGTATTT	GCAAAGCCCA	ATGCCATTCA	AGCGGTGGGG	420
GTATGTCTT	TTGCAATTT	TTGCCACCAT	AACTCCTTCT	TAGTTTACAG	TTCTCTAGAA	480
GAACCCACAG	TAGCTAAGTG	GTCCCGCCTT	ATCCATATGT	CCATCGTGAT	TTCTGTATTT	540
ATCTGTATAT	TCTTGTCTAC	ATGTGGATAC	TGACATTTA	CTGGCTTCAC	CCAAGGGGAC	600
TTATTTGAAA	ATTACTGCAG	AAATGATGAC	CTGGTAACAT	TTGGAAGATT	TTGTTATGGT	660
GTCACTGTCA	TTTGTACATA	CCCTATGGAA	TGCTTTGTGA	CAAGAGAGGT	AATTGCCAAT	720
GTGTTTTTTG	GTGGGAATCT	TTCAATCGTT	TTCCACATTG	TTGTAACAGT	GATGGTCATC	780
ACTGTAGCCA	CGCTTGTGTC	ATTGCTGATT	GATTGCCTCG	GGATAGTTCT	AGAAGTCAAT	840
GGTGTGCTCT	GTGCAACTCC	CCTCATTTTT	ATCATTCCAT	CAGCGTGTTA	TCTGAAAGTG	900
TCTGAAGAAC	CAAGGACACA	CTCCGATAAG	ATTATGTCTT	GTGTCATGCT	TCCCATTTGGT	960
GCTGTGGTGA	TGGTTTTTGG	ATTGCTCATG	GCTATTACAA	ATACTCAAGA	CTGCACCCAT	1020
GGGCAGGAAA	TGTTCTACTG	CTTTCCTGAC	AATTTCTCTC	TCACAAATAC	CTCAGAGTCT	1080
CATGTTTCAGC	AGACAACACA	ACTTTCTACT	TAAATATTA	GTATCTTTCA	ACTCGAGTAA	

Seq ID NO: 675 Protein sequence
Protein Accession #: Eos sequence

45

50

55

1	11	21	31	41	51	
MGYQRQEPVI	PPQVNTKTFP	PGYLLLSVLQ	FLYPFIAMIS	YNIIAGDTLS	KVFPQIPGVD	60
PENVFIGRHF	IIGLSTVTFP	LPLSLYRNIA	KLKGVSLIST	GLTTLILGIV	MARAILSGPH	120
IPKTEDAWVF	AKPNAIQAVG	VMSFAPICHH	NSFLVYSSLE	EPTVAKWSRL	IHMSIVISVP	180
ICIFFATCGY	LTFGTGTQGD	LFENYCRND	LVTFRGFCY	VTIVILTYPM	CFVTREIVAN	240
VFFGGNLSSV	FHVIVVTVMV	TVATLVSLLI	DCLGIVLELN	GVLCATPLIF	IIPSACYLKL	300
SEEPRTSDK	IMSCVMLPIG	AVVMVFGFVM	AITNTQDCTH	GQEMFYCFPD	NFSLTNTSES	360
HVQQTQLST	LNISIFQLE					

Seq ID NO: 676 DNA sequence
Nucleic Acid Accession #: NM_006853.1
Coding sequence: 26..874

60

65

70

75

80

1	11	21	31	41	51	
AGGAATCTGC	GCTCGGGTTC	CGCAGATGCA	GAGGTTGAGG	TGGCTGCGGG	ACTGGAAGTC	60
ATCGGGCAGA	GGTCTCACAG	CAGCCAAGGA	ACCTGGGGCC	CGCTCCTCCC	CCCTCCAGGC	120
CATGAGGATT	CTGCAGTTAA	TCCTGCTTGC	TCTGGCAACA	GGGCTGTAG	GGGGAGAGAC	180
CAGGATCATC	AAGGGGTTCC	AGTGCAAGCC	TCACTCCCAG	CCCTGGCAGG	CAGCCCTGTT	240
CGAGAAGACG	CGGCTACTCT	GTGGGGCGAC	GCTCATCGCC	CCCAGATGGC	TCCTGACAGC	300
AGCCCACTGC	CTCAAGCCCC	GCTACATAGT	TCACCTGGGG	CAGCACAACC	TCCAGAAGGA	360
GGAGGGCTGT	GAGCAGACCC	GGACAGCCAC	TGAGTCCTTC	CCCCACCCCG	GCTTCAACAA	420
CAGCCTCCCC	AACAAGACCC	ACCGCAATGA	CATCATGCTG	GTGAAGATGG	CATCGCCAGT	480
CTCCATCACC	TGGGCTGTGC	GACCCCTCAC	CCTCTCCTCA	CGCTGTGTCA	CTGCTGGCAC	540
CAGCTGCCCT	ATTTCGGGCT	GGGGCAGCAC	GTCCAGCCCC	CAGTTAAGCC	TGCCCTCACAC	600
CTTGOGATGC	GCCAACATCA	CCATCATTTGA	GCACCAGAAG	TGTGAGAACG	CCTACCCCGG	660
CAACATCACA	GACACCATGG	TGTGTGCCAG	CGTGACAGAA	GGGGGCAAGG	ACTCCTGCCA	720
GGGTGACTCC	GGGGGCCCTC	TGGTCTGTAA	CCAGTCTCTT	CAAGGCATTA	TCTCCTGGGG	780
CCAGGATCCG	TGTGCGATCA	CCCGAAAGCC	TGGTGTCTAC	ACGAAAGTCT	GCAAAATATGT	840
GGACTGAGTC	CAGGAGACGA	TGAAGAACAA	TTAGACTGGA	CCCCCCACCC	ACAGCCCATC	900
ACCTCCATT	TCCACTTGGT	GTITGGTTCC	TGTTCACTCT	GTTAATAAGA	AACCTTAAGC	960
CAAGACCTTC	TAGAACATT	CTTTGGGCTC	CCTGGACTAC	AGGAGATGCT	GTCACTTAAT	1020
AATCAACCTG	GGGTTCGAAA	TCAGTGAGAC	CTGGATTCAA	ATTCTGCCTT	GAAATATTTG	1080
GACTCTGGGA	ATGACAACAC	CTGTTTGTIT	CTCTGTGTGA	TCCCCAGCCC	CAAAGACAGC	1140
TCTTGCCCAT	ATATCAAGGT	TTCAATAAAT	ATTTGCTAAA	TGAGTG		

Seq ID NO: 677 Protein sequence
Protein Accession #: NP_006844.1

85

1	11	21	31	41	51	
MRILQLILLA	LATGLVGGET	RIIKGFCEPK	HSQPWQAALP	EKTRLLCGAT	LIAPRWLLTA	60

AHCLKPRYIV HLGQHNLOKE EGCEQTRTAT ESPPHPGFNN SLPNKDHRND IMLVKHASFV 120
SITWAVRPLT LSSRCVTAQT SCLISGWGST SSPQLRLPHT LRCANITIIIE HQKCNAYPG 180
NITDTMVCS VQEGGKDSQ GDSGGPLVCN QSLQGIISWG QDPCAITRKP GVTYKVCYV 240
DWIQETMKN

5

Seq ID NO: 678 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..933

10

1 11 21 31 41 51
ATGTGCAGCA ATGGACGGTG CATCCCGGGC GCCTGGCAGT GTGACGGGCT GCCTGACTGC 60
TTCGACAAGA GTGATGAGAA GGAGTGCCCC AAGGCTAAGT CGAAATGTGG CCGACCTTC 120
TTCCTCTGTG CCAGCGGCAT CCATTGCATC ATTGGTGTCT TCCGGTGCAA TGGGTTTGTG 180
GACTGTCCCG ATGGCAGCGA TGAAGAGAAC TGCACAGCAA ACCCTCTGCT TTGCTCCACC 240
GCCCGCTACC ACTGCAAGAA CGGCTCTGT ATTGACAAGA GCTTCATCTG CGATGGACAG 300
AATAACTGTG AAGACAACAG TGATGAGGAA AGCTGTGAAA GTTCTCAAGA ACCCGGCAGT 360
GGGCAGGTGT TTGTGACTCT AGAGAACCAG CTGTGTATT ACCCCAGCAT CACCTATGCC 420
ATCATCGGCA GCTCCGTCAT TTTGTGCTG GTGTGGGCC TGCTGGCACT GGTCTTGAC 480
CACCAGCGGA AGCGAACAAC CCTCATGACG CTGCCCGTGC ACCGCTGCA GCACCTGTG 540
CTGCTGTCCC GCCTGGTGGT CCTGGACAC CCCCACCACT GCAACGTGAC CTACAACGTC 600
AATAATGGCA TCCAGTAGT GGCACGCCAG GCGGAGCAGA ATGCGTCGGA AGTAGGCTCC 660
CCACCCTCT ACTCCGAGGC CTGTCTGGAC CAGAGGCGCT CGTGGTATGA CCTTCTCCA 720
CGCCCTACT CTCTGACAC GGAATCTCTG AACCAAGCGG ACCTGCCCCC CTACCGCTCC 780
CGGTCCGGGA GTGCCAACAG TGCCAGCTCC CAGGCAGCCA GCAGCTCTCT GAGGTGTGAA 840
GACACCGCC ACAGCCCGGG GCAGCTCTGC CCCCAGGAG GCACCTGTGA GCCCAGGGAC 900
TCTGAGCCA GCCAGGGCAC TGAAGAAGTA TAA

30

Seq ID NO: 679 Protein sequence
Protein Accession #: Eos sequence

1 11 21 31 41 51
MCSNGRCIPG ANQCDGLPDC FDKSDEKECP KAKSKCGPTF PPCASGIHCI IGRFRNGFE 60
DCPDGSDSEN CTANPLLCT ARYHCKNGLC IDKSPICDQ NNOQDNDSE SCSSQEPGS 120
GVVFTSENQ LVYPSITYA IIGSVIFVL VVALLALVLH HQRKRNLMT LPVHRLQHPV 180
LLSRLVLDH PHHCNVTYNV NNGIQVASQ AEQNASEVGS PPSYSEALD QRPAMYDLPP 240
PPYSSDTESL NQADLPYPYS RSGSANSASS QAASSLLSVE DTSHPGQPG PQBTAEPRD 300
SEPSQGTTEV

40

Seq ID NO: 680 DNA sequence
Nucleic Acid Accession #: S78203.1
Coding sequence: 1..2190

45 1 11 21 31 41 51
ATGAATCCTT TCCAGAAAA TGAGTCCAAG GAAACTCTTT TTTCACCTGT CTCCATTGAA 60
GAGGTACCAC CTGACCACC TAGCCCTCCA AAGAAGCCAT CTCGACAAT CTGTGGCTCC 120
AACTATCCAC TGAGCATTGC CTTCATTGTG GTGAATGAAT TCTGCGAGCG CTTTCTCTAT 180
ATTGGAATGA AAGCTGTGCT GATCCTGTAT TTCTGTATT TCCTGCACGT GAATGAAGAT 240
ACCTCCACAT CTATATAACA TGCCCTCAGC AGCCTCTGTT ATTTTACTCC CATCCTGGGA 300
GCAGCCATTG CTGACTCGTG GTTGGGAAAA TTCAAGACAA TCATCTATCT CTCCTGGTG 360
TATGTGCTTG GCCATGTGAT CAAGTCTCTG GGTGCCCTTAC CAATACTGGG AGGACAAGTG 420
GTACACACAG TCCATCATT GATCGGCTTG AGTCTAATAG CTTTGGGAGC AGSAGGCATC 480
AAACCCCTGT TGCAGCTTT TGGTGGAGAC CAGTTTGAAG AAAACATGC AGAGGAACGG 540
ACTAGATACT TCTCAGTCTT CTACCTGTCC ATCAATGCAG GGAGCTTGAT TTCTACATT 600
ATCACACCCA TGCTGAGAGG AGATGTGCAA TGTTTTGGAG AAGACTGCTA TGCATTGGCT 660
TTTGGAGTTC CAGGACTGCT CATGGTAATT GCACCTGTG TGTITGCAAT GGGAGCAAA 720
ATATACAATA AACCAACCCC TGAAGGAAAC ATAGTGGCTC AAGTTTTCAA ATGTATCTGG 780
TTTGTATTAT CCAATCGTTT CAAGAACCGT TCTGGAGACA TTCCAAAGCG ACAGCACTGG 840
CTAGACTGGG CAGCTGAGAA ATATCCAAG CAGCTCATTA TGGATGTAAG GGCACAGACC 900
AGGGTACTAT TCTTTTATAT CCCATTGCCC ATGTTCTGGG CTCTTTTGA TCAGCAGGT 960
TCACGATGGA CTTTGCAGC CATCAGGATG AATAGGAATT TGGGTTTGT TGTGCTTCAG 1020
CCGACCCAGA TGCAGGTCTT AAATCCCTTT CTGGTTCTTA TCTTCATCCC GTTGTTCAC 1080
TTTGTCTATT ATGCTCTGCT CTCCAAGTGT GGAATTAAT TCTCATCACT TAGGAAAATG 1140
GCTGTGCTGA TGATCCTAGC GTGCCTGGCA TTGTCAGTTG CGGCAGCTGT AGAGATAAAA 1200
ATAAATGAAA TGGCCCCAGC CCAATCAGGT CCCCAGGAGG TTTTCTACA AGTCTTGAAT 1260
CTGGCAGATG ATGAGGTGAA GGTGACAGTG GTGGGAAATG AAAACAATTC TCTGTTGATA 1320
GAGTCCATCA AATCCTTCA GAAACACCA CACTATTCCA AACTGCACCT GAAAACAAA 1380
AGCCAGGATT TTCATTCCA CCTGAAATAT CACAATTGT CTCTCTACAC TGAGCATTCT 1440
GTGCAGGAGA AGAAGTGTA CAGTCTGTG ATTOGTGAAG ATGGGAACAG TATCTCCAGC 1500
ATGATGGTAA AGGATACAGA AAGCAAAACA ACCAATGGGA TGACAACCGT GAGGTTTGT 1560
AACACTTTCG ATAAAGATGT CAACATCTCC CTGAGTACAG ATACCTCTCT CAATGTTGTT 1620
GAAGACTATG GTGTGCTGCT TTATAGAACT GTGCAAGAG GAGAATACCC TGCAGTGCAC 1680
TGTAAGACAG AAGATAAGAA CTTTCTCTG AATTTGGGTC TTCTAGACTT TGGTGCAGCA 1740
TATCTGTTTG TTATTACTAA TAACACCAAT CAGGGTCTTC AGGCTGGAA GATTGAAGAC 1800
ATTCCAGCCA ACAAAATGTC CATTGCGTGG CAGCTACCA ATATGCCCT GGTACAGCT 1860
GGGGAGGTCA TGTCTCTGT CACAGGTCTT GAGTTTCTT ATTCTCAGG TCCTCTAGC 1920
ATGAAATCTG TGCTCCAGGC AGCTTGGCTA TTGACAATTG CAGTTGGGAA TATCATCTG 1980
CTTGTGTGG CACAGTTCAG TGGCTGGTA CAGTGGGCGG AATTCATTT GTTTCTCTG 2040
CTCTGCTGG TGATCTGCCT GATCTCTCC ATCATGGCT ACTACTATGT TCCTGTAAAG 2100
ACAGAGGATA TGGGGGTCT AGCAGATAAG CACATTCCTC ACATCCAGG GAACATGATC 2160
AAACTAGAGA CCAAGAAGAC AAAACTCTGA

85

Seq ID NO: 681 Protein sequence
Protein Accession #: AAB34388.1

1 11 21 31 41 51
| | | | |
5 MNPFQKNBSK ETLFSPVSI E VPPRPSPSP KKPSPICGS NYPLSIAPIV VNEFCERFSY 60
YGMKAVLILY FLYFLHWNED TSTSIYHAFS SLCYFTPILG AALADSWLKG FKTIIYLSLV 120
YVLGHVIKSL GALPILGQOV VHTVLSLIGL SLIALGTGGI KPCVAAPGGD QPEEKHAER 180
TRYFSVYLLS INAGSLISTF ITPMLRGDVQ CPGEDCYALA FGVPGLLMVI ALVVVFAMGSK 240
IYNKPPPEGN IVAQVFKCIW FAISNRFKNR SGDIPKRQHW LDWAAEKYPK QLIMDVKALT 300
RVLFPLYLPL MFNALLDQOG SRWTLQAIRN NRNLGPPVLQ PDQMQLNPF LVLIFIPLED 360
FVIRLVSKC GINFSSLRJW AVGMILACLA FAVAAVEIK INEMAPAQSG PQEVFLQVLN 420
10 LADDEVKQTV VGNENNSLLI ESIKSFQKTP HYSKLHLKTK SQDFHFLKY HNLSTYTHS 480
VQKNWYSLV IREDGNSISS MMVKDTESKT TNGMTTVRFV NTLHKDVNIS LSTDTSLNVG 540
EDYGVSAIYR VQRGEYPAVH CRTEDKNFSL NLGLLDPGAA YLFVITNTN QGLQAWKIED 600
IPANKMSIWL QLPQYALVTA GEVMSVTGL EFSYSQAPSS MKSVLQAAML LTIAGNIIV 660
15 LVVAQPSGLV QWAEFILFSC LLLVICLIFS IMGYVVPVK TEDMRGPADK HIFHIQGNMI 720
KLETKKTKL

Seq ID NO: 682 DNA sequence
Nucleic Acid Accession #: NM_016077.1
Coding sequence: 128..657

20 1 11 21 31 41 51
| | | | |
25 TGGCTTTGTG ATTCTTGATC CGGAACCTTG TCACCCAGGA ACCCGGAAG AGGTAGCTCA 60
CGGATAGAA ACGTGTTCG TGGCCAGAA GAAGGGAAG CGGAGTGAG GAAAGGAGT 120
ACTGTAGATG CCTCCAAAT CCTGGTTAT GGAATATTG GCTCATCCA GTACTCTGG 180
CTGGCTGTT GGAGTTGCTT GTGGCATGTG CTGGGCTGG AGCCTTCGAG TATGCTTTGG 240
GATGCTCCC AAAAGCAAGA CGAGCAAGAC ACACACAGAT ACTGAAAGT AAGCAAGCAT 300
CTGGGAGAC AGCGGGGAGT ACAAGATGAT TCTGTGTT CGAATGACT TAAAGATGGG 360
30 AAAAGGAAA GTGGCTGCC AGTGCTCTCA TGCTGCTGT TCAGCTACA AGCAGATTCA 420
AAGAAGAAAT CCTGAAATG TCAACAATG GGAATACTG GGCAGCCCA AGTGGTGGT 480
CAAGCTCCT GATGAAGAAA CCTGATTGC ATTATTGGCC CATGCAAAA TGCTGGGACT 540
GACTGTAAGT TTAATTCAAG ATGCTGGACG TACTCAGAT GCACCAAGCT CTCAAAGTGT 600
CCTAGGATT GGGCAGGAC CAGCAGACT AATTGACAAA GTCACTGGTC ACCTAAAGT 660
35 TACTAGGTG GACTTGATA TGACAACAAC CCTCCATCA CAAGTGTGT AAGCCTGTCA 720
GATTCTAACA ACAAAAGCTG AATTTCTTCA CCCAAGTTAA ATGTTCTTGA GATGAAATA 780
AAACCTATTC CCATGTTCTA AAAAAA

Seq ID NO: 683 Protein sequence
Protein Accession #: NP_057161.1

40 1 11 21 31 41 51
| | | | |
45 MFSKSLVMEY LAHPSTLGLA VGVACGMCLG WSLRVCPGML PKSKTSKTHT DTESEASILG 60
DSGEYKMLV VRNDLKMKGK KVAAQCSHAA VSAYKQIQR NPMLKQWEY CGQPKVVVKA 120
PDEETLIALI ARAKMLGLTV SLIQDAGRTQ IAPGSQTVLG IGPSPADLID KVTGHLKLY

Seq ID NO: 684 DNA sequence
Nucleic Acid Accession #: NM_004864.1
Coding sequence: 26..952

50 1 11 21 31 41 51
| | | | |
55 CGGAACGAGG GCAACCTGCA CAGCCATGCC CGGGCAAGAA CTCAGGACGG TGAATGGCTC 60
TCAGATGCTC CTGGTGTTC TGGTGTCTC GTGGCTGCG CATGGGGCG CCTGTCTCT 120
GGCCGAGGCG AGCCGCGCAA GTTCCCGGG ACCCTCAGAG TTGCACTCG AAGACTCCAG 180
ATTCGAGAG TTGCGGAAG GCTACGAGGA CTGTCTAACC AGGCTCGGG CCAACAGAG 240
CTGGGAAGAT TCGAACACCG AACTCGTCCC GGCCTTGCA GTCCGATAC TCACGCCAG 300
AGTGGCGGTG GGATCCGCG GCCACCTGCA CTRGCGTATC TCTCGGGCG CCTTCCCGA 360
GGGGCTCCCC GAGGCTCCC GCCTTCAACG GGCTCTGTT CGGCTGTCCC CGACGGGTC 420
60 AAGGTCTGTG GACGTGACAC GACCGCTGCG GCGTCAGCTC AGCCTTGCAA GACCCCAAGC 480
GCCCGCGCTG CACTGCGAC TGTGCGCCG GCGTGCAG TCGGACCAAC TGCTGGCAGA 540
ATCTTCGTCC GCACGCCCC AGCTGGAGTT GCACCTGCG CGCAAGCCG CCAGGGGGCG 600
CGCAGAGCG CGTGCAGCA ACGGGGACGA CTGTCCGCTC GGGCCCGGGC GTTGCTGCGC 660
TCTGCACAG GTCCGCGGT CGCTGGAAGA CTTGGGCTGG GCGGATTGG TGCTGTGCGC 720
65 ACGGGAGGTG CAAGTGACCA TGTGATCGG CGGCTGCGG AGCCAGTTCC GGGCGGCAA 780
CATGCAGCG CAGATCAAGA CGAGCTGCA CGCCTGAAG CCGACACGG AGCCAGCGCC 840
CTGCTGCGT CCCGCCAGT ACAATCCCAT GGTGCTCATT CAAAGACCG ACACCGGGT 900
GTGCTCCAG ACCTATGAT ACTGTTAGC CAAAGACTGC CACTGCATAT GAGCAGTCT 960
GGTCTCTCCA CTGTGCACT GCGCGGGGA GCGGACCTCA GTTGTCTGCT CCTGTGGAAT 1020
70 GGGCTCAAG TTCTGAGAC ACCCGATTCC TGCCCAACA GCTGTATTTA TATAAGTCTG 1080
TTATTATTA TTAATTTATT GGGGTGACCT TCTTGGGAC TCGGGGGCTG GTCTGATGGA 1140
ACTGTGTATT TATTTAAAC TCTGTGATA AAAATAAAGC TGCTGAACT GTTAAAAAAA 1200
AAAA

Seq ID NO: 685 Protein sequence
Protein Accession #: NP_004855.1

75 1 11 21 31 41 51
| | | | |
80 MPQQLRLTVN GSQMLLVLLV LSWLPHGGAL SLAEASRAS PGPSELHSED SRFRELRKRY 60
EDLLTRLRAN QSWEDSNTDL VPAPAVRILT PEVRLSGGGH LHLRISRAAL PEGLPASRL 120
HRLRLSLPT ASRSNDVTRP LRRQLSLARP QAPALHLRLS PPPSQSDQLL AESSSARPOL 180
ELHLRPQAR GRRRARARNG DDCLPGPRC RLHTVRAEL EDLGWADWVL SPREVQVTC 240
85 IGACPSQFRA ANMHAQIKTS LHRLKPDTEP APCCVPASYN PMVLIQKTDV GVSLLQTYDDL 300
LAKDCHCI

Seq ID NO: 686 DNA sequence

Nucleic Acid Accession #: NM_002423.2
Coding sequence: 48..851

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      ACCAAATCAA CCATAGGTCC AAGAACAATT GTCTCTGGAC GGCAGCTATG CGACTCACCG 60
      TGCTGTGTGC TGTGTGCTCT CTGCCTGGCA GCCTGGCCCT GCGCTGCTCT CAGGAGGCGG 120
      GAGGCATGAG TGAGCTACAG TGGGAACAGG CTCAGGACTA TCTCAAGAGA TTTTATCTCT 180
      ATGACTCAGA AACAAAAAAT GCCAACAGTT TAGAAGCCAA ACTCAAGGAG ATGCAAAAAT 240
10     TCTTTGGCCT ACCTATAACT GGAATGTTAA ACTCCCGCGT CATAGAAATA ATGCAGAAAGC 300
      CCAGATGTGG AGTGCCAGAT GTTGCAGAAT ACTCACTATT TCCAAATAGC CCAAAATGGA 360
      CTTCCAAAGT GGTCACTTAC AGGATCGTAT CATATACTCG AGACTTACCG CATATTACAG 420
      TGGATCGATT AGTGTCAAAG GCTTTAAACA TGTGGGGCAA AGAGATCCCC CTGCATTTCa 480
      GGAAGTTGTG ATGGGGAACT GCTGACATCA TGATTGGCTT TCGCGGAGGA GCTCATGGGG 540
15     ACTCTACACC ATTTGATGGG CCAGGAAACA CGCTGGCTCA TGCCCTTTGG CCTGGGACAG 600
      GTCTCGGAGG AGATGCTCAC TTCCGATGAG ATGAACGCTG GACGGATGGT AGCAGTCTAG 660
      GGATTAACCT CCGTATGCTG GCAACTCATG AACTTGGCCA TTCTTTGGGT ATGGGACATT 720
      CCTCTGATCC TAATGCAGTG ATGTATCCAA CCTATGGAAA TGGAGATCCC CAAATTTTA 780
      AACTTTCCCA GGATGATATT AAAGGCATTG AGAACTATA TGGAAAGAGA AGTAATTCAA 840
20     GAAAGAAATA GAAACTTCAG GCAGAACATC CATTCATTCA TTCATTGGAT TGTATATCAT 900
      TGTTCACAAA TCAGAAATGA TAAGCACTGT TCCTCCACTC CATTTAGCAA TTATGTCAAC 960
      CTTTTTATTG CGAGTGTGTT TTTGAATGTC TTTCACTCCT TTTATTGGTT AAACCTCCTT 1020
      ATGGTGTGAC TGTGTCTTAT TCCATCTATG AGCTTTGTCA GTGCGCGTAG ATGTCAATAA 1080
25     ATGTTACATA CACAAATAAA TAAATGTTT ATTCCATGGT AAATTTA

```

Seq ID NO: 687 Protein sequence
Protein Accession #: NP_002414.1

```

30     1      11      21      31      41      51
      |      |      |      |      |      |
      MRLTVLCAVC LLPGLSLALPL POEAGGMSLE QWEQAQDYLK RPYLYDSETH NANSLEAKLK 60
      EMQKFFGLPI TGLMNSRVIE IMQKPRCGVP DVAEYSLFPN SPKWTISKVVT YRIVSYTRDL 120
      PHITVDRLVS KALNMWKEIPI PLHFRKVVWG TADIMIGFAR GAHGDSYFPD GPQNTLAHAF 180
      APGTGLGGDA HFDDEDERWD GSSLGINFLY AATHELGHSL GMGHSSDPNA VMYPTYGNNGD 240
35     PQNFKLSQDD IKGIQKLYGK RSNRKK

```

Seq ID NO: 688 DNA sequence
Nucleic Acid Accession #: NM_005221.3
Coding sequence: 1..870

```

40     1      11      21      31      41      51
      |      |      |      |      |      |
      ATGACAGGAG TGTITGACAG AAGGGTCCCC AGCATCCGAT CCGGCGACTT CCAAGCTCCG 60
      TTCCAGACGT CCGCAGCTAT GCACCATCCG TCTCAGGAAT CGCCAACTTT GCCCGAGTCT 120
      TCAGCTACCG ATTCTGACTA CTACAGCCCT ACGGGGGGAG CCGCGCACGG CTACTGCTCT 180
      CTAACCTCGG CTCTCTATGG CAAAGCTCTC AACCCCTACC AGTATCAGTA TCACGGGCTG 240
      AACGGCTCCG CCGGAGCTA CCCAGCCAAA GCTTATGCCG ACTATAGCTA CGCTAGCTCC 300
      TACCACCACT ACGGCGGCGC CTACAACCGC GTCCCAAGCG CCACCAACCA GCCAGAGAAA 360
      GAAGTGACCG AGCCCGAGGT GAGAATGGTG AATGGCAAAC CAAAGAAAAGT TGTAAACCC 420
50     AGGACTATT ATTCCAGCTT TCAGCTGGCC GCATTACAGA GAAGGTTTCA GAAGACTCAG 480
      TACCTCGCCT TGCCGGAACG CGCCGAGCTG GCCGCTCGC TGGGATTGAC ACAAACACAG 540
      GTGAAATCT GGTTCAGAA CAAAAGATCC AAGATCAAGA AGATCATGAA AAACGGGGAG 600
      ATGCCCCCGG AGCACAGTCC CAGCTCCAGC GACCCATGG CGTGTAACCT GCGCGAGTCT 660
      CCAGCGGTGT GGGAGCCCCA GGGCTCGTCC CGCTCGCTCA GCCACCAACC TCATGCCCCA 720
55     CCTCCGAOCT CCAACCACTC CCCAGCGTCC AGCTACCTGG AGAACTCTGC ATCCTGGTAC 780
      ACAAGTGCAG CCAGCTCAAT CAATTCCAC CTGCCGCGC CGGGCTCCTT ACAGCACCCG 840
      CTGGCGCTGG CCTCCGGGAC ACTCTATTAG

```

Seq ID NO: 689 Protein sequence
Protein Accession #: NP_005212.1

```

60     1      11      21      31      41      51
      |      |      |      |      |      |
      MTGVFDRRVP SIRSGDFQAP FQTSAAHEHP SQESPTLPES SATOSDYISP TGGAPHGYCS 60
      PTSASYGKAL NPVQYQYHGV NGSAGSYPAK AYADYSYASS YHQYGGAYNR VPSATNQPEK 120
      EVTEPEVRMV NGKPKKVRKP RTIYSSPQLA ALQRRFQKTQ YLALPERAEL AASLGLTQTQ 180
      VKIWFQNKRS KIKKIMNGE MPPEHSPSSS DFMACNSPQS PAVWEPOGSS RSLSHHPAH 240
      PPTSQSPAS SYLENSASWY TSAASSINSH LPPPGSLQHP LALASGTLV

```

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

- 1 1. A method of detecting a lung cancer-associated transcript in a cell
2 from a patient, the method comprising contacting a biological sample from the patient with a
3 polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence
4 as shown in Tables 1A-16.
- 1 2. The method of claim 1, wherein the polynucleotide selectively
2 hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16.
- 1 3. The method of claim 1, wherein the biological sample is a tissue
2 sample.
- 1 4. The method of claim 1, wherein the biological sample comprises
2 isolated nucleic acids.
- 1 5. The method of claim 4, wherein the nucleic acids are mRNA.
- 1 6. The method of claim 4, further comprising the step of amplifying
2 nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1 7. The method of claim 1, wherein the polynucleotide comprises a
2 sequence as shown in Tables 1A-16.
- 1 8. The method of claim 1, wherein the polynucleotide is labeled.
- 1 9. The method of claim 8, wherein the label is a fluorescent label.
- 1 10. The method of claim 1, wherein the polynucleotide is immobilized on
2 a solid surface.
- 1 11. The method of claim 1, wherein the patient is undergoing a therapeutic
2 regimen to treat lung cancer.
- 1 12. The method of claim 1, wherein the patient is suspected of having lung
2 cancer.
- 1 13. A method of monitoring the efficacy of a therapeutic treatment of lung
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a lung cancer-associated transcript in the
6 biological sample by contacting the biological sample with a polynucleotide that selectively
7 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16,
8 thereby monitoring the efficacy of the therapy.

1 14. The method of claim 13, further comprising the step of: (iii) comparing
2 the level of the lung cancer-associated transcript to a level of the lung cancer-associated
3 transcript in a biological sample from the patient prior to, or earlier in, the therapeutic
4 treatment.

1 15. The method of claim 13, wherein the patient is a human.

1 16. A method of monitoring the efficacy of a therapeutic treatment of lung
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a lung cancer-associated antibody in the biological
6 sample by contacting the biological sample with a polypeptide encoded by a polynucleotide
7 that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in
8 Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated
9 antibody, thereby monitoring the efficacy of the therapy.

1 17. The method of claim 16, further comprising the step of: (iii) comparing
2 the level of the lung cancer-associated antibody to a level of the lung cancer-associated
3 antibody in a biological sample from the patient prior to, or earlier in, the therapeutic
4 treatment.

1 18. The method of claim 16, wherein the patient is a human.

1 19. A method of monitoring the efficacy of a therapeutic treatment of lung
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a lung cancer-associated polypeptide in the
6 biological sample by contacting the biological sample with an antibody, wherein the antibody
7 specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to
8 a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby
9 monitoring the efficacy of the therapy.

1 20. The method of claim 19, further comprising the step of: (iii) comparing
2 the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated
3 polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic
4 treatment.

1 21. The method of claim 19, wherein the patient is a human.

1 22. An isolated nucleic acid molecule consisting of a polynucleotide
2 sequence as shown in Tables 1A-16.

1 23. The nucleic acid molecule of claim 22, which is labeled.

1 24. The nucleic acid of claim 23, wherein the label is a fluorescent label

1 25. An expression vector comprising the nucleic acid of claim 22.

1 26. A host cell comprising the expression vector of claim 25.

1 27. An isolated polypeptide which is encoded by a nucleic acid molecule
2 having polynucleotide sequence as shown in Tables 1A-16.

1 28. An antibody that specifically binds a polypeptide of claim 27.

1 29. The antibody of claim 28, further conjugated to an effector component.

1 30. The antibody of claim 29, wherein the effector component is a
2 fluorescent label.

1 31. The antibody of claim 29, wherein the effector component is a
2 radioisotope or a cytotoxic chemical.

1 32. The antibody of claim 29, which is an antibody fragment.

- 1 33. The antibody of claim 29, which is a humanized antibody
- 1 34. A method of detecting a lung cancer cell in a biological sample from a
2 patient, the method comprising contacting the biological sample with an antibody of claim
3 28.
- 1 35. The method of claim 34, wherein the antibody is further conjugated to
2 an effector component.
- 1 36. The method of claim 35, wherein the effector component is a
2 fluorescent label.
- 1 37. A method of detecting antibodies specific to lung cancer in a patient,
2 the method comprising contacting a biological sample from the patient with a polypeptide
3 encoded by a nucleic acid comprises a sequence from Tables 1A-16.
- 1 38. A method for identifying a compound that modulates a lung cancer-
2 associated polypeptide; the method comprising the steps of:
3 (i) contacting the compound with a lung cancer-associated polypeptide, the
4 polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least
5 80% identical to a sequence as shown in Tables 1A-16; and
6 (ii) determining the functional effect of the compound upon the polypeptide.
- 1 39. The method of claim 38, wherein the functional effect is a physical
2 effect.
- 1 40. The method of claim 38, wherein the functional effect is a chemical
2 effect.
- 1 41. The method of claim 38, wherein the polypeptide is expressed in a
2 eukaryotic host cell or cell membrane.
- 1 42. The method of claim 38, wherein the functional effect is determined by
2 measuring ligand binding to the polypeptide.
- 1 43. The method of claim 38, wherein the polypeptide is recombinant.

1 44. A method of inhibiting proliferation of a lung cancer-associated cell to
2 treat lung cancer in a patient, the method comprising the step of administering to the subject a
3 therapeutically effective amount of a compound identified using the method of claim 38.

1 45. The method of claim 44, wherein the compound is an antibody.

1 46. The method of claim 45, wherein the patient is a human.

1 47. A drug screening assay comprising the steps of
2 (i) administering a test compound to a mammal having lung cancer or a cell
3 isolated therefrom;
4 (ii) comparing the level of gene expression of a polynucleotide that selectively
5 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16 in a
6 treated cell or mammal with the level of gene expression of the polynucleotide in a control
7 cell or mammal, wherein a test compound that modulates the level of expression of the
8 polynucleotide is a candidate for the treatment of lung cancer.

1 48. The assay of claim 47, wherein the control is a mammal with lung
2 cancer or a cell therefrom that has not been treated with the test compound.

1 49. The assay of claim 47, wherein the control is a normal cell or mammal.

1 50. A method for treating a mammal having lung cancer comprising
2 administering a compound identified by the assay of claim 47.

1 51. A pharmaceutical composition for treating a mammal having lung
2 cancer, the composition comprising a compound identified by the assay of claim 47 and a
3 physiologically acceptable excipient.

REVISED VERSION

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 2002/086443 A2

(51) International Patent Classification⁷: **C07H 21/02**,
21/04, C12Q 1/68

(21) International Application Number:
PCT/US2002/012476

(22) International Filing Date: 18 April 2002 (18.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/284,770 18 April 2001 (18.04.2001) US
60/290,492 10 May 2001 (10.05.2001) US
60/339,245 9 November 2001 (09.11.2001) US
60/350,666 13 November 2001 (13.11.2001) US
60/334,370 29 November 2001 (29.11.2001) US
60/372,246 12 April 2002 (12.04.2002) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— of inventorship (Rule 4.17(iv)) for US only

(71) Applicant (*for all designated States except US*): PRO-TEIN DESIGN LABS, INC [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080 (US).

Published:
— with declaration under Article 17(2)(a); without abstract; title not checked by the International Searching Authority

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): AZIZ, Natasha [US/US]; 411 California Avenue, Palo Alto, CA 94306 (US). MURRAY, Richard [US/US]; 22643 Woodbridge Court, Cupertino, CA 95014 (US).

(48) Date of publication of this revised version: 17 June 2004

(15) Information about Correction:
see PCT Gazette No. 25/2004 of 17 June 2004, Section II

(74) Agent: HALLUIN, Albert, P.; HOWREY SIMON ARNOLD & WHITE LLP, 301 Ravenswood Avenue, Box 34, Menlo Park, CA 94025 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

(57) Abstract:

WO 2002/086443 A2

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rule 13ter.1(c) and 39)

Applicant's or agent's file reference 18501-15-3PC	IMPORTANT DECLARATION	Date of mailing (day/month/year) 15 AUG 2003
International application No. PCT/US02/12476	International filing date (day/month/year) 18 April 2002 (18.04.2002)	(Earliest) Priority date (day/month/year) 10 May 2001 (10.05.2001)
International Patent Classification (IPC) or both national classification and IPC IPC(7): C07H 21/02, 21/04; C12Q 1/68 and US Cl.: 435/6, 536/23.1, 23.5		
Applicant EOS BIOTECHNOLOGY, INC		

This International Searching Authority hereby declares, according to Article 17(2)(a), that no international search report will be established on the international application for the reasons indicated below.

1. ☐ The subject matter of the international application relates to:
 - a. ☐ scientific theories.
 - b. ☐ mathematical theories
 - c. ☐ plant varieties.
 - d. ☐ animal varieties.
 - e. ☐ essential biological processes for the production of plants and animals, other than microbiological processes and the products of such processes.
 - f. ☐ schemes, rules or methods of doing business.
 - g. ☐ schemes, rules or methods of performing purely mental acts.
 - h. ☐ schemes, rules or methods of playing games.
 - i. ☐ methods for treatment of the human body by surgery or therapy.
 - j. ☐ methods for treatment of the animal body by surgery or therapy.
 - k. ☐ diagnostic methods practised on the human or animal body.
 - l. ☐ mere presentations of information.
 - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art.
2. ☐ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:

☐ the description
☐ the claims
☐ the drawings
3. ☒ The failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions prevents a meaningful search from being carried out:

☐ the written form has not been furnished or does not comply with the standard.
☒ the computer readable form has not been furnished or does not comply with the standard.
4. Further comments:

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer <div style="text-align: center;"> CARLA J. MYERS PRIMARY EXAMINER </div> Telephone No. 703-308-0196
--	---

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
TOWNSEND AND TOWNSEND AND CREW LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT OR THE DECLARATION

(PCT Rule 44.1)

Applicant's or agent's file reference 18501-15-3PC	Date of Mailing (day/month/year) 15 AUG 2003
International application No. PCT/US92/12476	International filing date (day/month/year) 18 April 2002 (18.04.2002)
Applicant EOS BIOTECHNOLOGY, INC	

1. ☐ The applicant is hereby notified that the international search report has been established and is transmitted herewith.

Filing of amendments and statement under Article 19:
The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

Where? Directly to the International Bureau of WIPO, 34, chemin des Colombettes
1211 Geneva 20, Switzerland, Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☒ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. Reminders

Shortly after 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90 *bis*.1 and 90 *bis*.3, respectively, before the completion of the technical preparations for international publication.

Within 19 months from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later); otherwise the applicant must, within 20 months from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

In respect of other designated Offices, the time limit of 30 months (or later) will apply even if no demand is filed within 19 months.

See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer <i>Valerie Bell-Harris</i> Carla Myers Telephone No. 703-308-0196
--	---

Form PCT/ISA/220 (April 2002)

(See notes on accompanying sheet)